



## **Novel Approaches to Cancer Therapy**

1. Allosteric inhibitors of the Ras-activating exchange factor Sos
2. Inhibition of interferon regulatory factor 4 (IRF4) with drug-like-small-molecules as a useful therapy in lymphoid cancers
3. FragLites: Developing new approaches to assess druggability and hit finding
4. MDM2-PROTACS for the Androgen Receptors (ARs) as a prostate cancer treatment

---

**Jose Daniel Lopez Fernandez**

Supervisors:  
Dr. Ian R. Hardcastle  
Professor Michael J. Waring

<b>ABSTRACT</b>	<b>8</b>
<b>CHAPTER 1: NOVEL APPROACHES TO CANCER THERAPY</b>	<b>17</b>
1.1 Cancer	17
1.2 Cancer Impact	18
1.3 Hallmarks of Cancer	21
1.4 Cancer Chemotherapy	24
1.5 Drug Discovery nowadays	27
1.5.1 The development of a Drug Discovery project	27
1.5.2 Tool Compounds or Chemical Probes	28
1.6 Fragment-based drug discovery (FBDD)	29
1.7 Biophysical Assays for validation of hits and lead compounds	29
<b>CHAPTER 2: INTRODUCTION TO ALLOSTERIC INHIBITORS OF THE RAS-ACTIVATING EXCHANGE FACTOR SOS</b>	<b>30</b>
2.1 A Ras History	31
2.2 Ras Functions	39
2.3 The structure of the G Domain	40
2.3.1 The Structural Changes of the switch	41
2.3.2 Moving out the nucleotide, GEF reaction	41
2.3.3 Son of Sevenless (Sos)	43
2.4 Effectors	44
2.5 GAP proteins and GTPase Reaction	44
2.6 Is Ras Druggable?	46
2.6.1 Targeting Ras directly	46
2.6.2 New and alternative approaches to target Ras	56
2.7 Future Perspectives	59
<b>CHAPTER 3: PROJECT RATIONALE</b>	<b>60</b>
3.1 Research Plan	66
3.1.1 Optimization of compound (11) and (12)	66
<b>CHAPTER 4: RESULTS AND DISCUSSION</b>	<b>69</b>
<b>CHAPTER 5: BIOLOGICAL ASSAYS</b>	<b>82</b>
<b>CHAPTER 6: CONCLUSIONS AND FUTURE WORK</b>	<b>88</b>

<b>CHAPTER 7: INTRODUCTION TO INHIBITION OF IRF4 WITH DRUG-LIKE SMALL MOLECULES AS A USEFUL THERAPY IN LYMPHOID CANCERS</b>	<b>89</b>
7.1 Transcription Factors and cancer	89
7.1.1 IRF family	89
7.1.2 IRF-4 and cancer	91
7.2 Transcription factors as drug targets	95
7.3 Current Approaches to inhibit IRF4 for multiple myeloma malignancies	96
7.3.1 Lenalidomide	96
7.3.2 Bortezomib combined with lenalidomide	97
<b>CHAPTER 8: PROJECT RATIONALE</b>	<b>100</b>
8.1 Structural Biology and Hit Discovery	103
<b>CHAPTER 9: RESULTS AND DISCUSSION</b>	<b>110</b>
9.1 Methodology	110
9.2 Targets synthesized	110
9.2.1 Pyridazinones	110
9.2.2 Acrylamides	118
<b>CHAPTER 10: BIOPHYSICAL ASSAYS AND CONCLUSIONS AFTER HTS AND SAR FOCUSED ON THE DNA BINDING DOMAIN (DBD)</b>	<b>121</b>
<b>CHAPTER 11: FRAGMENT-BASED SCREENING BY SPR FOCUSED ON THE DNA BINDING DOMAIN (DBD)</b>	<b>124</b>
<b>CHAPTER 12: DISCUSSION AND RESULTS OF THE NEW CHEMICAL SERIES</b>	<b>128</b>
<b>CHAPTER 13: BIOPHYSICAL ASSAYS AFTER FBS AND SAR FOCUSED ON THE DNA BINDING DOMAIN (DBD)</b>	<b>134</b>
<b>CHAPTER 14: CONCLUSIONS AND FUTURE WORK</b>	<b>141</b>
<b>CHAPTER 15: FRAGLITES. DEVELOPING NEW APPROACHES TO ASSESS DRUGGABILITY AND HIT FINDING AT EARLIER STAGES</b>	<b>143</b>
15.1 Improvements in the scientific methodology to speed-up the drug discovery process	143
<b>CHAPTER 16: PROJECT RATIONALE</b>	<b>150</b>
<b>CHAPTER 17: RESULTS AND DISCUSSION</b>	<b>154</b>

<b>CHAPTER 18: CONCLUSION AND FUTURE WORK</b>	<b>161</b>
<b>CHAPTER 19: INTRODUCTION TO MDM2 PROTACS FOR THE ANDROGEN RECEPTORS AS A PROSTATE CANCER TREATMENT</b>	<b>162</b>
19.1 <b>PROTACs: Inducing protein degradation</b>	<b>162</b>
19.1.1 Background	162
19.1.2 Mechanism of action	164
19.2 <b>Alternative approaches to protein degradation</b>	<b>167</b>
19.3 <b>Androgen Receptors and PROTACs</b>	<b>186</b>
19.3.1 The case of androgen receptor (AR) and flutamide	186
19.4 <b>MDM2 as an E3 ligase ligand</b>	<b>189</b>
19.4.1 MDM2 biology	189
19.4.2 MDM2 inhibitors	190
<b>CHAPTER 20: PROJECT RATIONALE</b>	<b>191</b>
20.1 <b>Design of a hetero-functional molecule</b>	<b>191</b>
20.1.1 E3 ligase ligand	191
20.1.2 Androgen Receptor Inhibitor	194
20.1.3 Linker exploration	194
<b>CHAPTER 21: RESULTS AND DISCUSSION</b>	<b>195</b>
21.1 <b>Methodology and Synthesis</b>	<b>195</b>
21.1.1 Part A: flutamide	195
21.1.2 Part B: E3 ligase ligand	196
21.1.3 Part C: linker	197
21.1.4 Tethering of the different parts	198
<b>CHAPTER 22: BIOLOGICAL ASSAYS</b>	<b>207</b>
22.1 <b>Western blot analysis (enzalutimide as control)</b>	<b>209</b>
22.1.1 Addition of PROTACs LNCaP cells	209
22.1.2 Addition of PROTACs to VCaP cells:	210
22.1.3 Addition of PROTACs to CWR22Rv1 cells:	211
22.1.4 Addition of PROTACs to AR-Vs in AR-EK cells:	212
22.2 <b>Western blot analysis (flutamide as control)</b>	<b>212</b>
22.2.1 Addition of PROTACs LNCaP cells:	213
22.2.2 Addition of PROTACs to VCaP cells:	213
22.3 <b>Testing the hook effect</b>	<b>214</b>
22.3.1 Addition of enzalutamide in LNCaP and VCaP cells	214
22.3.2 PROTAC 5 in LNCaP and VCaP cells:	215
22.3.3 PROTAC 2 in LNCaP and VCaP cells:	216
22.3.4 Addition of PROTAC 2 and Nutlin-3a to LNCaP cells:	217
22.3.5 Addition of PROTAC 2 and Nutlin-3a to VCaP cells	217
22.4 <b>Cell proliferation assays</b>	<b>218</b>
22.4.1 PROTAC 1 cell proliferation assay	218



22.4.2	PROTAC 2 cell proliferation assay	218
22.4.3	PROTAC 3 cell proliferation assay	219
22.4.4	PROTAC 4 cell proliferation assay	219
22.4.5	PROTAC 5 cell proliferation assay	220
<b>CHAPTER 23: CONCLUSIONS AND FUTURE WORK</b>		<b>221</b>
<b>CHAPTER 24: EXPERIMENTAL AND ANALYSIS</b>		<b>224</b>
24.1	<b>Safety</b>	<b>224</b>
24.2	<b>Solvents and Reagents</b>	<b>224</b>
24.3	<b>Chromatography and Equipment</b>	<b>224</b>
24.4	<b>Analytical Techniques</b>	<b>224</b>
<b>CHAPTER 25: GENERAL EXPERIMENTAL</b>		<b>226</b>
25.1	<b>Synthesis of (15) and (17) and derivatives:</b>	<b>226</b>
25.2	<b>Synthesis of (54) and derivatives:</b>	<b>244</b>
25.3	<b>Synthesis of (80) and derivatives</b>	<b>256</b>
25.4	<b>Synthesis of (85) and derivatives</b>	<b>259</b>
25.5	<b>Synthesis of (86) and derivatives</b>	<b>265</b>
25.6	<b>Synthesis of FragLites derivatives</b>	<b>271</b>
25.7	<b>Synthesis of flutamide precursor</b>	<b>275</b>
25.8	<b>Synthesis of (132)</b>	<b>276</b>
25.9	<b>Synthesis of MDM2-based PROTACs for ARs</b>	<b>279</b>

## **Declaration**

The work contributing to this thesis was conducted between October 2015 and September 2019 in the Medicinal Chemistry laboratories, Bedson Building, Northern Institute for Cancer Research, Newcastle University, Newcastle upon Tyne, NE1 7RU.

All the research presented in this thesis is original in context and does not include any material or ideas previously published or presented by other authors except where due reference is given in the text. Any figures included without reference were generated using ChemDraw or Microsoft Office and were drawn from knowledge.

No part of this thesis has been previously submitted for a degree, diploma or any other qualification at any other university

## Acknowledgements

I would like to thank my primary PhD supervisor, Ian R. Hardcastle for his support and encouragement over the past four years. He has been an excellent mentor both academically and emotionally and I am eternally grateful for his guidance during my PhD. I would also like to thank my other supervisors, Prof. Mike J. Waring, Prof. Bernard Golding and Dr. Celine Cano for their invaluable knowledge and expertise and continued enthusiasm during my project. I would like to thank all the people I have been privileged to work with in the lab at the Bedson Building and at the Paul O’Gorman Building. I would also like to thank Julie Tucker and Duncan C. Miller, who were a continual support and provided invaluable biological and chemical knowledge during my thesis.

I would also like to thank co-workers at the NICR for their continued help with the bioscience, Prof. Martin Noble and Prof. Steve Wedge for all their help with structural biology and Anita Wittner and Daniel J. Wood for conducting cell-based assays.

Finally, I would like to thank Dr. Leonelo Forti Sampietro for his guidance and support, my close friends and family, particularly my mother for her support throughout my thesis and all my life. Also, to Anxela for always being there to help when things were not going as good as we would like, and for always believing in me. **I cannot thank you enough.**

## ABSTRACT

### Allosteric inhibitors of the Ras-activating exchange factor SoS

SoS protein is responsible for keeping Ras, a key molecular switch involved in cell division, turned on. In many cancers a mutation in the Ras protein can permanently lock the molecular switch into the 'on' position, leading to enhanced cell division. This problem is especially relevant to pancreatic cancer, which has a particularly high prevalence of mutant Ras driven tumours (90%). Inhibiting the ability of SoS to catalyse nucleotide exchange on Ras has the potential to afford a means of targeting oncogenic Ras signalling.

In preliminary studies, we have identified small molecule compounds that bind to an unprecedented site on SoS itself and have the potential to inhibit Ras activation via a novel allosteric mechanism. Currently, we are exploring and modifying the structure of this small molecule to achieve a more potent and drug-like compound.

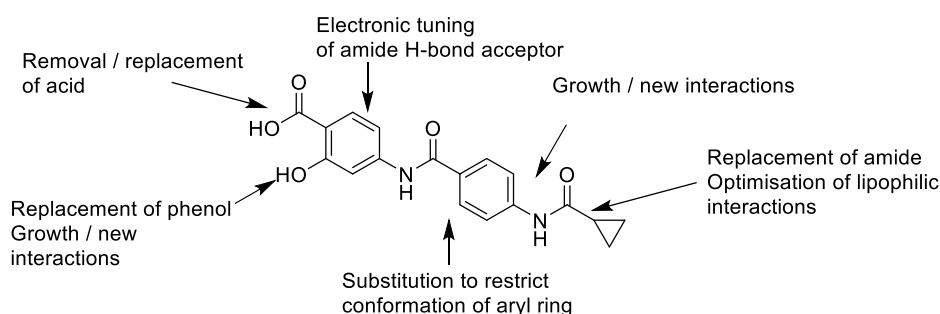


Figure A1; Outline chemical strategy for chemical optimization

### Small-molecule inhibitors of interferon regulatory factor 4 (IRF4)

Interferon regulatory factor 4 (IRF4) is a transcription factor that has been shown to regulate the survival of several aggressive lymphoid malignancies. It appears to be a master-regulator of gene transcription in myeloma, influencing cellular survival, proliferation, metabolism and differentiation. siRNA knockdown of IRF4 by only 40 – 60% is lethal to myeloma cells. The possibility of inhibiting IRF4 with drug-like small-molecules could be a useful therapy in a number of aggressive lymphoid cancers.

High Throughput Screen using Fluorescence Polarization assay for the IRF4 DNA Binding Domain and a fragment-based screen using SPR and NMR were undertaken. 7 hits were

obtained from both screens and 63 close analogues were synthesised to explore SAR and address physicochemical properties.

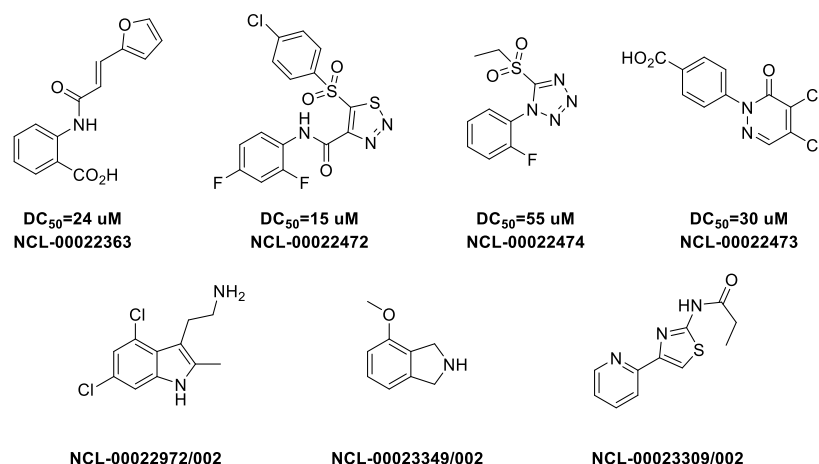


Figure A2; Hits obtained after HTS and FBS

### FragLites: Bifunctional Halogenated Fragments

Identification of ligand binding sites on proteins is of critical importance for chemical biology and drug discovery applications. We describe an efficient and comprehensive approach to mapping interaction sites on a protein using a designed set of bifunctional halogenated fragments that express hydrogen-bonding pharmacophore doublets. These fragments are capable of seeking out productive polar interactions and can be identified unambiguously by X-ray crystallography by the anomalous scattering of their halogen substituent.

Soaking this set of compounds with CDK2 elucidated previously identified binding sites and also new ones, which can be used as starting points for fragment growing and drug design.

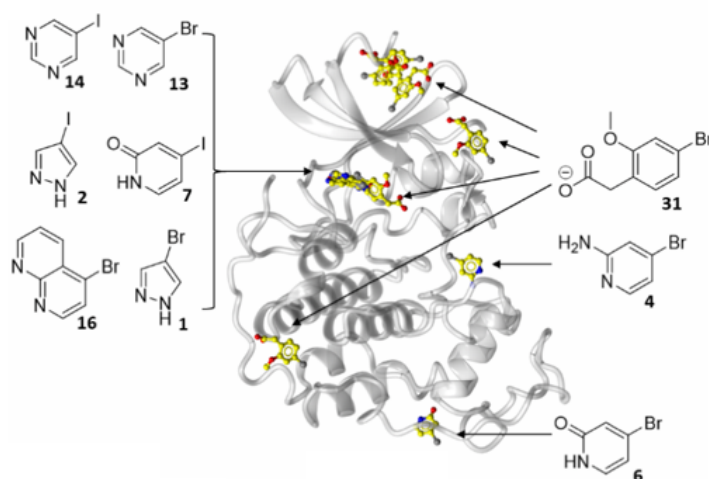


Figure A3; Overview of Fraglite binding events to CDK2

### MDM2-PROTACS for the Androgen Receptors (AR)

PROTACS (proteolysis targeting chimeric molecules) have been recently disclosed as chemical tools to target specific proteins for destruction by the proteasome. PROTACS comprise a ligand for an E3 ligase protein that may be a peptide or small molecule, and a ligand for the target protein. The advent of highly potent and specific ligands for MDM2 (E3 ligase) offers the opportunity to develop MDM2-PROTACS. Several compounds were designed to target the Androgen Receptor via Flutamide using two different MDM2 ligands and a broad variety of linkers (Figure 4).

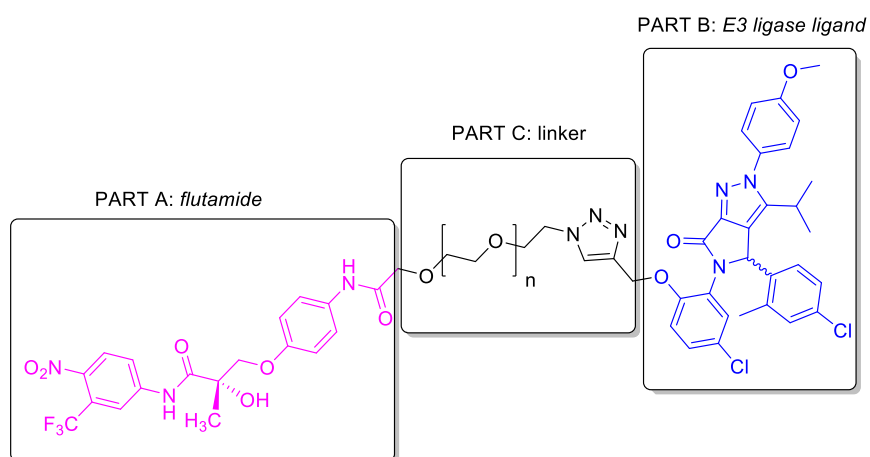


Figure A4; PROTAC molecule divided by functionalities (in pink: AR inhibitor, in blue: MDM2 inhibitor, in black: linker)

We have demonstrated the activity of MDM2 PROTACS for degradation of the AR and confirmed the importance of the linker length in different *in vitro* assays. The influence of the potency of the E3 ligase ligand was also confirmed by SPR experiments.

## Abbreviations

AAAS	American Association for the Advancement of Science
ADMET	Absorption, Distribution, Metabolism, Elimination and Toxicology
ADT	Androgen Deprivation Therapy
AR	Androgen Receptors
ARVs	Androgen Receptor Variants
ATP	Adenosine Tri-Phosphate
B-CLL	B-Cell Chronic Lymphocytic Leukaemia
BET	Bromo- and Extra- Terminal domain
Brd4	Bromodomain 4
$\beta$ -TRCP	$\beta$ -Transducing Repeat-Containing Protein
CDK2	CycloDextrin Kinase 2
CRABPI	Cellular Retinoic Acid Binding Protein I
CRBN	Cereblon
CRL	Cullin-Ring Ligase
CRPC	Castration Resistant Prostate Cancer
CRUK	Cancer Research UK
DABCO	1,4-Di-Aza-Bi-Cyclo[2.2.2]Octane
DBD	DNA Binding Domain
DCM	Dichloro Methane
DDR	DNA Damage Response
DHP	DiHydroPyran
DIPEA	Di-Isopropyl Ethyl Amine
DMAP	Di-Methyl Amino Pyridine
DMSO	DiMethyl Sulfoxide

DMF	Di-Methyl Formamide
DNA	Deoxyribonucleic acid
DSF	Differential Scanning Fluorimetry
DHT	Di-Hydro-Testosterone
E26	E26 Transformation-Specific
EDCI	1-Ethyl-3-(3-Dimethylaminopropyl)Carbodi-Imide
ER	Estrogen Receptor
ERK	Extra-cellular signal-Regulated Kinase
FBDD	Fragment-Based Drug Discovery
FBS	Fragment-Based Screening
FDA	Food and Drug Administration
FP	Fluorescence Polarisation
FS	Fragment Screening
FTIs	Farnesyl Transferase Inhibitors
GAPs	GTPase Activating Proteins
GDP	Guanosine Di-Phosphate
GEF	Guanine Nucleotide Exchange Factor
GNBPs	Guanine Nucleotide Binding Proteins
GPCRs	G-Protein Coupled Receptors
GST	Glutathione S-Transferase
GTP	Guanosine Tri-Phosphate
HIF1 $\alpha$	Hypoxia-Inducible Factor 1 $\alpha$
H2L	Hit to Lead
HATU	Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium
HBA	Hydrogen Bond Acceptor



HBD	Hydrogen Bond Donnor
HBTU	Hexafluorophosphate Benzotriazole Tetramethyl Uronium
HEPES	4-(2-Hydroxy-Ethyl)-1-Piperazineethanesulfonic Acid
HIV	Human Immunodeficiency Virus
HyT	Hydrophobic Tagging
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
HRS	Hodgkin and Reed Sternberg
HSP	Heat-Shock Protein
HTRF	Homogeneous Time-Resolved Fluorescence
HTS	High-Throughput Screening
IAD	Interferon Association Domain
IAP	Inhibitor of Apoptosis Protein
ICMT	Isoprenylcysteine Carboxyl Methyl Transferase
ICS	Interferon Consensus Sequence
IFN	InterFeroN
IL-6	InterLeukin-6
IPTG	Iso-Propyl $\beta$ -D-1-Thio-Galactopyranoside
IRF	Interferon Regulatory Factor
ISRE	Interferon Stimulated Response Element
ITC	Isothermal Titration Calorimetry
IZF	Ikaros family Zinc Finger
K <sub>d</sub>	Constant of dissociation
LE	Ligand Efficiency
MAP-2	Methionine Amino Peptidase-2

MAPKs	Mitogen-Activated Protein Kinases
MCL	Mantle Cell Lymphoma
MDM2	Murine Double-Minute 2
MM	Multiple Myeloma
mM	miliMolar
MoA	Mechanism of Action
mPCa	metastatic Prostate Cancer cells
MST	Micro-Scale Thermophoresis
mTOR	mammalian Target Of Rapamycin
MW	Molecular Weight
NICR	Northern Institute for Cancer Research
nM	nanoMolar
NNRTI	<i>Non-Nucleoside Reverse Transcriptase</i> Inhibitor
NTA	Nitrilo Triacetic Acid
PanDDA	Pan-Dataset Density Analysis
PDAC	Pancreatic Ductal AdenoCarcinoma
PDB ID	Protein Data-Base Identification
PI3K	Phosphoinositide-3-Kinase
POI	Protein Of Interest
PPIs	Protein-Protein Interactions
pRb	protein Retinoblastoma-associated
PROTACs	Proteolysis Targeting Chimeras
PSA	Polar Surface Area
RBD	Ras-Binding Domain
RCE1	Ras Converting Enzyme 1

REM	Ras Exchanger Motif
RNA	RiboNucleic Acid
ROS	Reactive Oxygen Species
r.t.	room temperature
RTKs	Receptor Tyrosine Kinases
SAR	Structure-Activity-Relationship
SARD	Selective Androgen Receptor Degradors
SAXS	Small Angle X-ray Scattering
SERD	Selective Estrogen Receptor Degradors
SERM	Selective Estrogen Receptor Modulators
SKP1	S-phase Kinase-associated Protein 1
Sos	Son of Sevenless
SPR	Surface Plasmon Resonance
STAT	Signal Transducer and Activator of Transcription Element
STD	Saturation-Transfer Difference
TAMRA	Tetramethyl-Rhodamine Azide
TB	Tuberculosis
TBL	Tracheal, Bronchus and Lung
TFA	Tri-Fluoroacetic Acid
THF	TetraHydroFuran
THP	TetraHydroPyrane
T <sub>m</sub>	Thermal Melting
TROSY	Transverse Relaxation Optimized Spectroscopy
UPS	Ubiquitin-Proteasome Pathway
VEGF	Vascular endothelial growth factor

VHL Von-Hippel Lindau

$\mu\text{M}$  micromolar

## Chapter 1: Novel approaches to cancer therapy

### 1.1 Cancer

Cancer is a group of complex genetic diseases characterised by alterations in normal cell growth and function resulting in uncontrolled proliferation.<sup>1</sup> Over 100 types of cancer and over 200 subtypes of tumours affect humans in different organs and tissues.<sup>2</sup> Most of these cancers are named by the organ or cell type where they originated.

Cancer malignancies are caused by alterations in the DNA sequence after several genetic (or epigenetic) mutations happen in a somatic cell. These can lead to stimulation of oncogenes, which instead of regulation will cause proliferation and subsequent survival of tumour cells. In many cases the loss (or under-expression) of suppressor genes involved in the inhibition of proliferation and apoptosis will happen in parallel.<sup>3</sup> Exogenous factors can accelerate or promote the appearance of cancer (*e.g.*: life-style and environmental factors such as alcohol abuse; smoking or sedentary routines are known to induce cancer). Other risk factors to take into account are endogenous factors, which may involve inherited genetic mutations and hormones (Figure 1). It is also known that age is the biggest risk factor and should be considered when we analyse the prevalence of the disease in developed societies.<sup>4</sup> Half of all cancer cases in the UK each year are diagnosed in people aged  $\geq 70$  years.

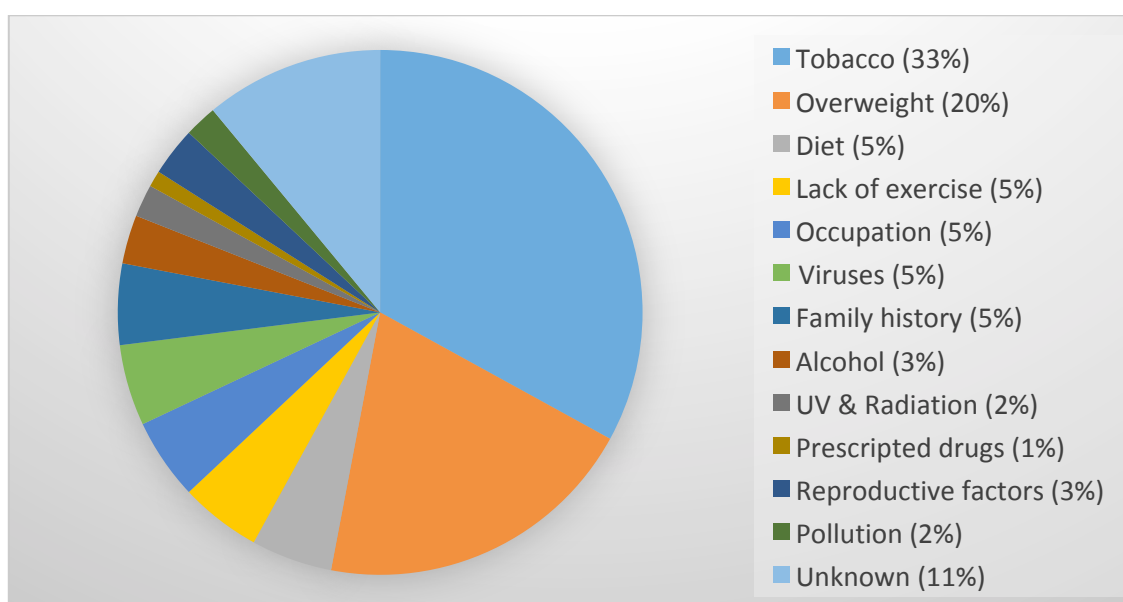


Figure 1; Common causes associated with cancer in USA<sup>5</sup>

## 1.2 Cancer Impact

Cancer has an immense impact on public health across the world. Since 1990, the world has experienced an increase in life expectancy and therefore the incidence of cancer and its mortality has increased similarly. Lifestyles and behaviours associated with cancer including smoking, physical inactivity, and diet are widespread in developing countries and the incidence and mortality rates show significant regional variability (Figure 2).

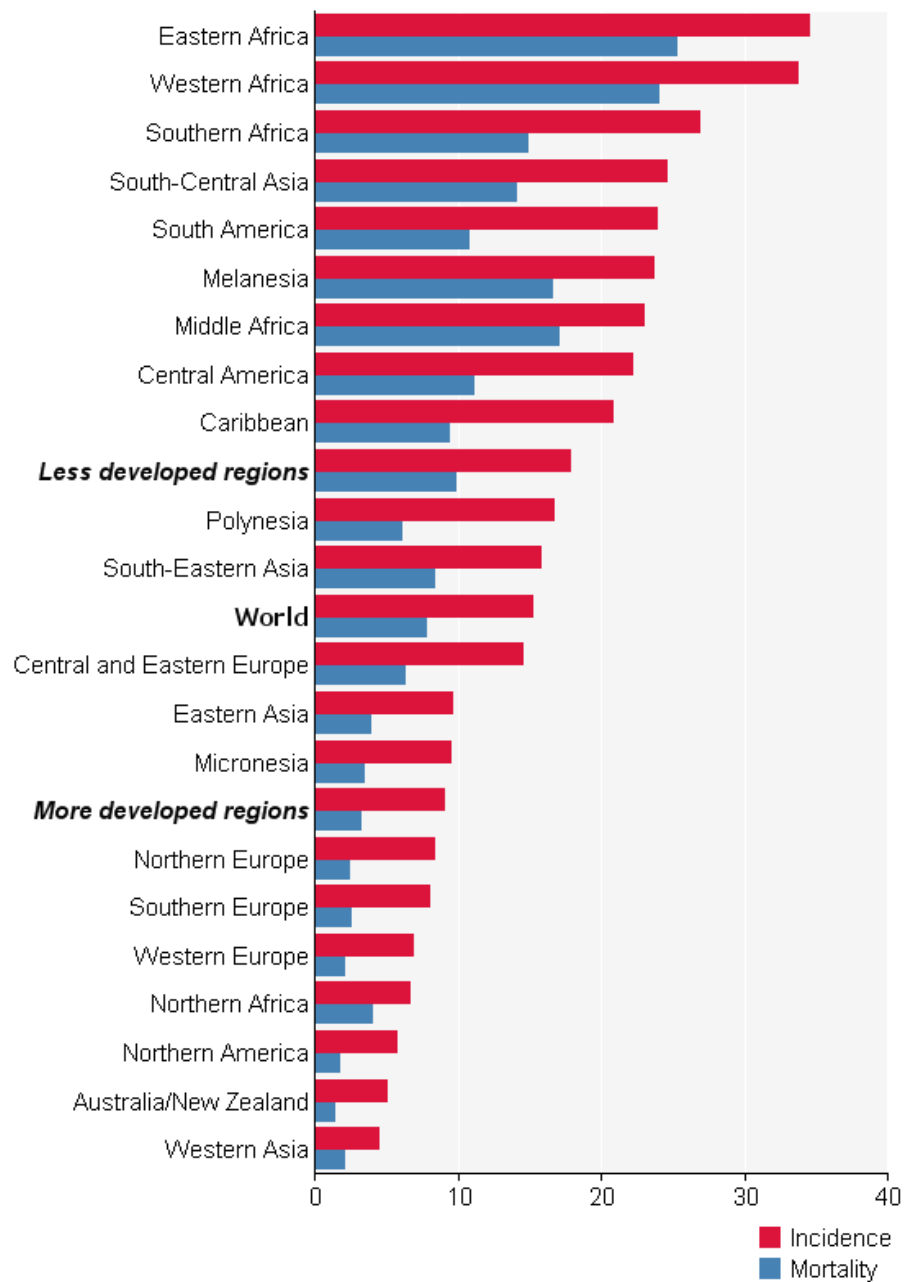


Figure 2; Age-standardized incidence and mortality rates in 2018 (per 100,000)<sup>6</sup>

Cancer has become the second leading cause of death in the world following heart disease. An increase of 33% in cancer cases over the previous decade has provoked an estimated of 17.5 million new cancer cases and 8.7 million deaths in 2015. Of all the different cancer types, tracheal, bronchus and lung (TBL) cancers are the most prevalent of the list with 1.6 million deaths. In women, breast cancer is the most frequently reported (28% of total incidence), and is the leading cause of cancer death. In men, prostate cancer is the most commonly reported, comprising 17% of total new cancer cases with 1.6 million cases.

Since the 1980s, a steady increase has been observed in the number of cancer cases, and according to a study conducted by Cancer Research UK (CRUK) in 2015, half of British people will develop cancer at some point in their lives. In 2015, breast, prostate, lung and bowel cancers contributed more than 50% of all cases (Figure 3).<sup>7</sup>

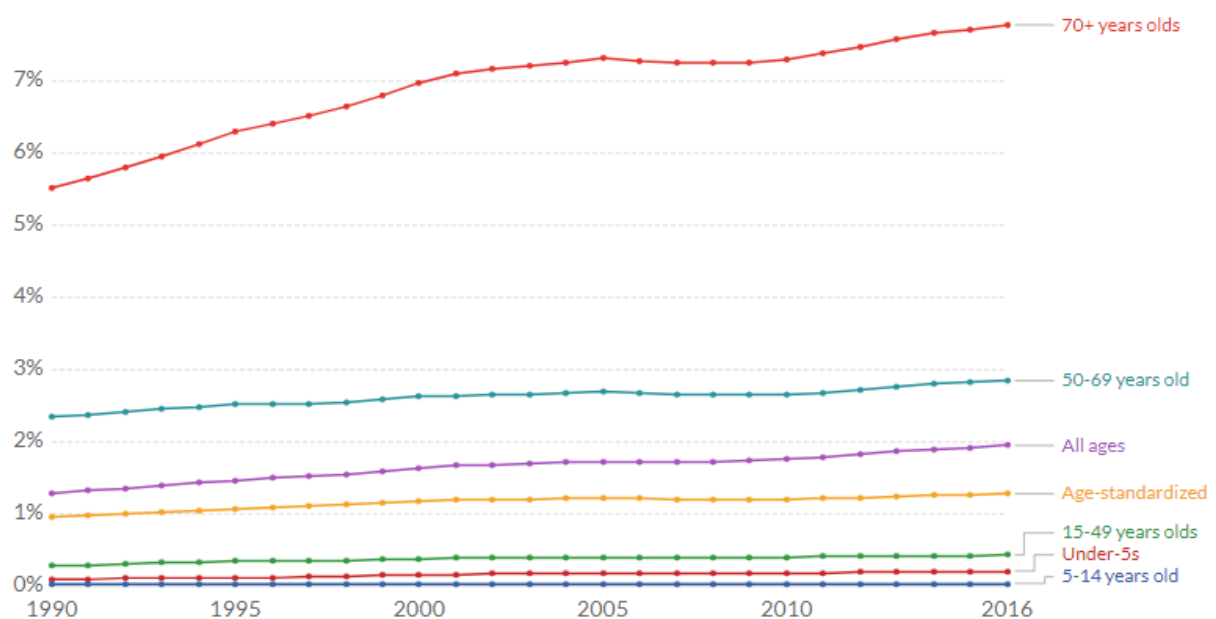


Figure 3, Share of population with cancer by age in the UK

However, there are reasons for optimism; the ten-year survival rate has doubled since the 1980s. In addition, overall cancer death rates in the UK have declined by 15% since 1990 (Figure 4).

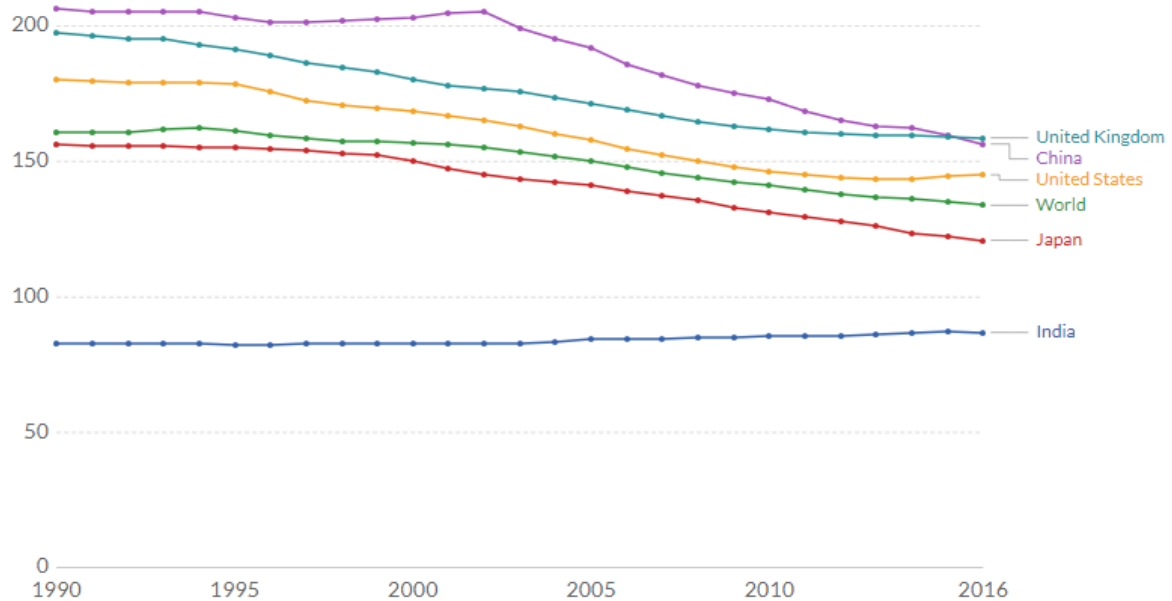


Figure 4; Cancer death rates (per 100000) age-standardized from all cancer types<sup>8</sup>

Advances in our understanding of cancer and the processes through which cells progress into a tumor state have yielded therapeutic drugs for the disease. Analysing the 5-year survival rate over the last 4 decades (indicating the percentage of people who lived longer than five years following diagnosis) we can find cases where life expectancy has doubled (Figure 5).

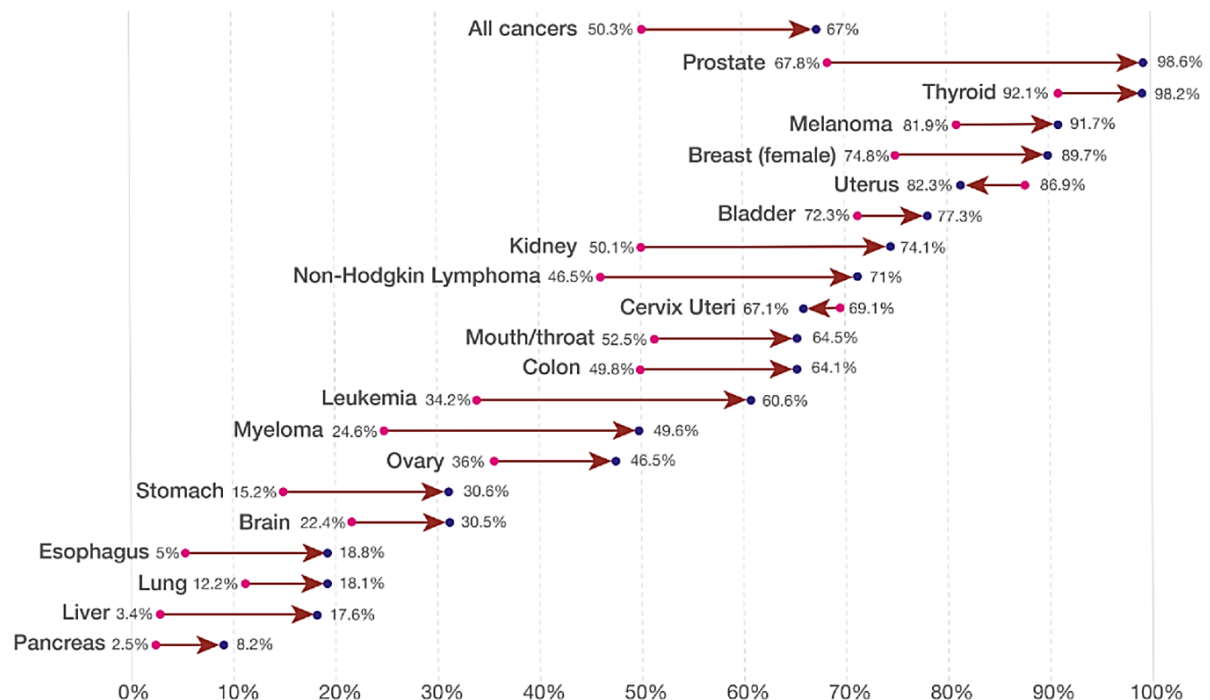


Figure 5; Average five-years survival rates from common cancer types, 1970-77 to 2007-13



Despite these significant improvements there are still a number of specific cancers, such as lung and pancreatic cancers, for which not enough advances have been made. More research is needed to achieve new therapeutic agents.

### 1.3 Hallmarks of Cancer

During the past 30 years, the complicated mechanisms by which cancer is driven have been deeply studied and the conclusions published in a great research effort. We know that tumorigenesis is a multi-step pathway and that these steps reflect several genetic mutations exquisitely synchronized that drive the conversion of normal cells into cancerous cells.

In 2000, Hanahan and Weinberg published a comprehensive review in which they described six acquired capabilities or physiological changes known as the hallmarks of cancer.<sup>9</sup> The number was increased to ten in a more recent review.<sup>10</sup> Most tumour types in humans share the distinct features presented in these studies, allowing them to survive, proliferate and spread (Figure 6).

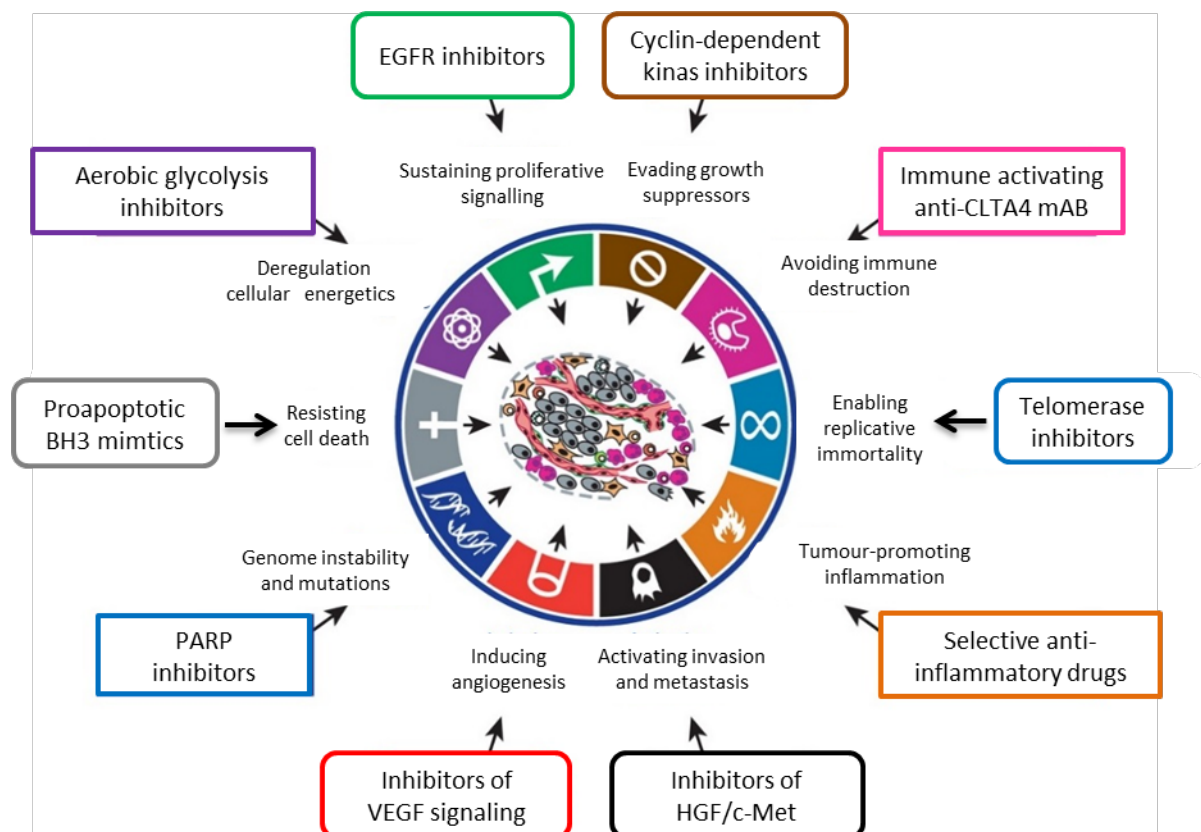


Figure 6; Hallmarks of cancer<sup>12</sup>

This study provided an important overview of the basic physiology of a mutant cell and the progressive transformations that they induce in their immediate environment. These hallmarks are described below:

### ***Sustaining proliferative signalling***

Cell growth and proliferation is controlled by growth factors that pass the information inside the cell via trans-membrane receptors. These growth factors can be diffusible, extracellular matrix components or cell-to-cell adhesion/interaction molecules. In the case of cancer cells, there is a developed independence from exogenous signals. Cancer cells can proliferate through alternative pathways (upstream or downstream) like creating their own growth factors or becoming hyper-responsive to the environment via over-expression of epidermal growth factor receptor. Structural mutations in the receptor proteins can also induce hyper-responsiveness, which contribute to growth and proliferation of cancer cells without exogenous control.

### ***Evading growth suppressors***

Quiescence and normal tissue homeostasis keep the balance between growth factors and antiproliferation factors. These factors include several tumour suppressor genes such as pRb (retinoblastoma-associated) or p53, which limit cell growth and division. These genes are frequently found mutated in many types of cancer cancers, resulting in inactivation of antiproliferation mechanisms and therefore, cancer cells can evade the antigrowth signals.

### ***Resisting cell death***

There are two main apoptotic pathways: the intrinsic and the extrinsic. In the intrinsic pathway cells respond to internal signals (DNA damage or oncogene expressed proteins), while external signals (hormones or cytokines) activate the extrinsic pathway. Proteases involved in these pathways can lead to destruction of cells by phagocytes. There are several mechanisms by which cancer cells resist apoptosis, one of the most studied is the over-expression of proteins involved in suppressing apoptosis.

### ***Enabling replicative immortality***

Before going into a non-proliferative state (senescence) cells have a limited number of cycles of cell division. This is determined by the telomeres, which are essential to protect the

chromosomes against instability through the prevention of chromosomal degradation. The length of telomeres decreases with each phase of DNA replication until it becomes completely degraded, triggering a crisis mode and causing apoptosis. *Telomerase* is a DNA polymerase that extends telomeric DNA, enhancing the replicative potential of cells. Tumour cells acquire unlimited replicative potential due to over-expression of the *telomerase* enzyme.

### ***Inducing angiogenesis***

Tumour cells have the ability to establish their own vascular system in a process called angiogenesis. Thanks to this tumour, cells can obtain nutrients and oxygen. Vascular endothelial growth factor (VEGF) is the most potent factor that initiates angiogenesis in the surroundings of the tumour and within.

### ***Invasion and metastasis***

Metastasis is the ability of tumour cells to invade neighbouring tissues and spread through the body via the lymphatic and blood circulatory system. E-cadherin is a protein that promotes the adhesion of an individual cell to other cells in the surrounding area and is down regulated in tumours, which allows metastatic processes to happen by the development of secondary malignant growths that are distant from the primary site of cancer.

### ***Genome instability and mutations***

To maintain genome integrity and to counteract DNA damage, cells have developed DNA-damage response (DDR) pathways. These pathways allow cells to detect DNA damage and activate cell cycle arrest for DNA damage repair. If the damage cannot be repaired, the DDR pathways will lead the cell to apoptosis and prevent the propagation of the damage. Malfunctioning DDR pathways lead to damaged DNA replication, which leads to genome instability and causes progression to cancer. The inactivation of tumour suppressor genes can occur through epigenetic mechanisms such as DNA methylation and histone modifications.

### ***Tumour-promoting inflammation***

Tumour-associated inflammatory states are driven by cells of the immune system. Some of which serve to promote tumorigenesis. Inflammatory cells can release reactive species (*e.g.* reactive oxygen radicals (ROS)), which cause DNA damage and can lead to genomic

instability. Inflammatory cells also provide bioactive molecules such as growth factors, resulting in sustained proliferative signalling.

### ***Deregulation of cellular energetics***

The modification of cellular energy metabolism is necessary to support sustained neoplastic growth and proliferation. Normal cells utilise glucose to produce energy through glycolysis in the presence of oxygen thanks to the mitochondria. In the absence of oxygen, normal cells rely on anaerobic glycolysis. Regardless of the availability of oxygen, cancer cells can upregulate their glucose metabolism by increasing aerobic glycolysis. Mutations affecting enzymes related to metabolic processes (*e.g.* p53) are known to reprogram cancer metabolism as well.

### ***Avoiding immune destruction***

The immune system can recognise and eliminate emerging tumours. In order to escape from attack and elimination by immune cells, cancer cells can be enabled with several mechanisms providing this ability. For example, secreting immunosuppressive factors, which prevent natural killer cells from eliminating tumour cells.

These Hallmarks are cellular mechanisms by which cancer cells differ from a normal cell. From a therapeutic point of view, all the different strategies followed by cancer cells offer us an opportunity to disrupt the tumorigenic process via **Cancer Chemotherapy**.

## **1.4 Cancer Chemotherapy**

The traditional treatments for cancer are radiotherapy or surgery. In the case of surgical treatment, removing the tumour could solve the problem, whereas by the application of X-rays we try to kill cancer cells and shrink the tumour. The success of both approaches depends on early diagnosis, in general are not successful against invasive cancers. The introduction of chemical compounds as a viable therapeutic approach has come to be known as **chemotherapy**. This strategy started in the 1950s with nitrogen mustard<sup>11</sup> (Figure 7-compound 1), to treat non-Hodgkin's lymphoma. In the following years, new DNA-alkylating agents were introduced, including orally administered drugs like chlorambucil<sup>12</sup> (Figure 7-compound 2).

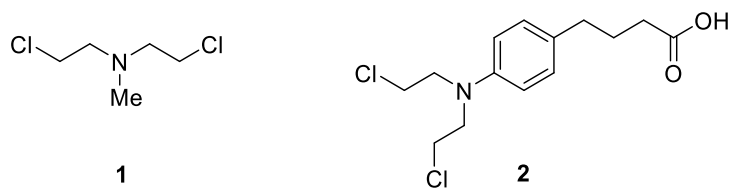


Figure 7; Nitrogen mustard derivative (**1**) and Chlorambucil (**2**)

In 1965, the combination chemotherapy was hypothesized by James F. Holland, Emil Freireich and Emil Frei by following the same principles of antibiotic therapy. A multidrug approach is followed in most cases where different mechanisms of actions can work together and overcome resistance mechanisms or bypass pathways that help cancer cells to survive.<sup>13</sup>

### ***Cytotoxic cancer therapy***

In general, chemotherapeutic drugs and macromolecules cannot target proteins in an entirely selective process, leading to cytotoxicity and cell death. Cancer cells use normal cells in their neoplastic activity, and a huge number of normal cells are directly involved in the replicating process that can be affected by chemotherapy. The damage provoked in these cells result in side-effects that in general lead to nausea, vomiting and hair loss. Furthermore, cytotoxic treatment also leads to drug resistance as result of new mutations in the tumour cells. These new mutations are the most common reason for treatment failure. The often necessary large doses to cover all organs and achieve a therapeutic efficacy against tumour cells cause more side-effects, and faster acquired resistance has led to new strategies directing drug accumulation in cancerous cells. One of these strategies involves designing compounds that exploit processes in tumour physiology that are not present in healthy cells.

Despite the limitations of traditional cytotoxic chemotherapy, it is still the treatment of choice for most types of cancer and the search to find new chemotherapeutics is still ongoing (Figure 8).<sup>14</sup>

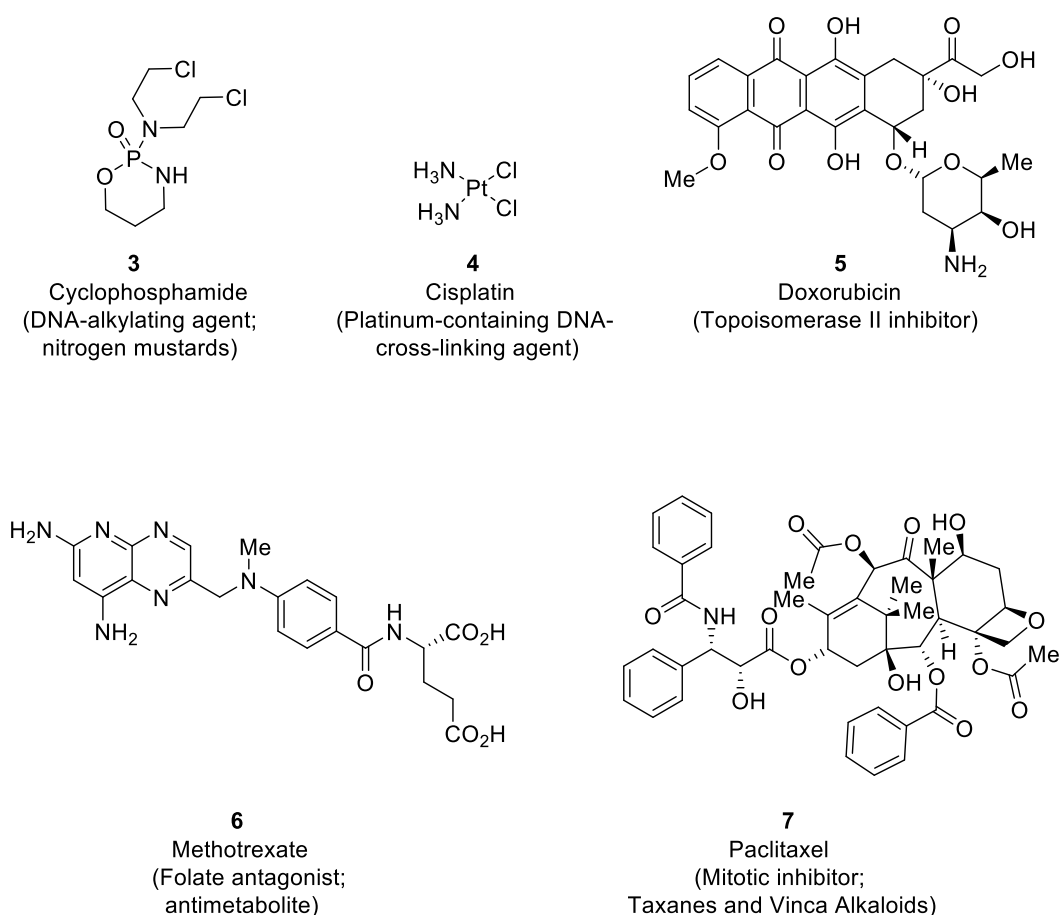


Figure 8; Common cytotoxic agents used in the treatment of cancer

## Targeted cancer chemotherapy

In the past twenty years we have seen great developments in the field of chemotherapy. Thanks to a better understanding of cancer progress and evasion of our immune system, new targeted cancer therapies like monoclonal antibodies and small molecule inhibitors are emerging. By targeting growth factors, signalling molecules, cell-cycle proteins and processes involved in apoptosis and angiogenesis, which are crucial for tumourigenesis, we can be more specific between cancer cells and healthy cells. Also, targeted therapies are administered at the optimal effective dose rather than at the maximum tolerated dose, which avoids high toxicity and side effects.<sup>15</sup>

## 1.5 Drug Discovery nowadays

### 1.5.1 The development of a Drug Discovery project

Drug discovery is a long and expensive process. It takes an average of 12–15 years and about \$1 billion to produce a new drug (Figure 9).<sup>16</sup>



Figure 9; Drug discovery stages

The initial step is to identify a target and validate it, so we can treat a specific protein or a biochemical pathway to affect the disease. This step can be approached by bioinformatics, biochemical *in vivo* and *in vitro* experimental models, genetic interaction and functional screening, including knockdown or using chemical probes. Once the biological target is validated, the next stage is to identify hit molecules with activity against the target. This process can include high-throughput screening (HTS), where a library of thousands of compounds is screened, fragment based-screening (FBS), virtual screening, or structure-based design.

After hit identification, the hit-2-lead (H2L) stage is carried out by running systematic structure-activity relationship (SAR) studies. The purpose is to determine the core structure (pharmacophore) and the modifications that will affect the potency and selectivity of the compound. If successful, a lead compound will be obtained. Structure-based drug design strategies using molecular modelling and technologies like X-ray crystallography and NMR will guide medicinal chemists to develop the SAR rationally.

Lead compounds are further optimised during the lead optimisation campaign using a range of *in vitro* and *in vivo* ADMET (absorption, distribution, metabolism, elimination, toxicology) assays to improve the properties of the candidate drug.

Finally, one or a few compounds are progressed to pre-clinical stage to assess *in vivo* toxicology before being advanced to clinical trials.

### 1.5.2 Tool Compounds or Chemical Probes

Tool compounds, also known as probes, are potent, selective and cell-permeable small molecules that play an essential role in the target validation stage. The genomic era has yielded an increasing number of potential druggable targets (an estimated 5,000-10,000) with a possible therapeutic effect. Chemical tools are necessary to validate a specific strategy by clarifying the relationship between a therapeutic target and the biological response. These tools allow researchers to explore if a specific target modulation by small molecules can lead to a novel drug.

Different biophysical and biochemical techniques can successfully identify a specific chemical probe or tool compound. However, the successful application of these probes remains challenging, and we can often find several examples in the literature where chemical probes have been misused. Only high quality chemical probes generate meaningful biological data. It is difficult for the research community to keep well informed about the most appropriate chemical probe and some initiatives have appeared like a Chemical Probe Portal (Figure 10).<sup>17</sup>

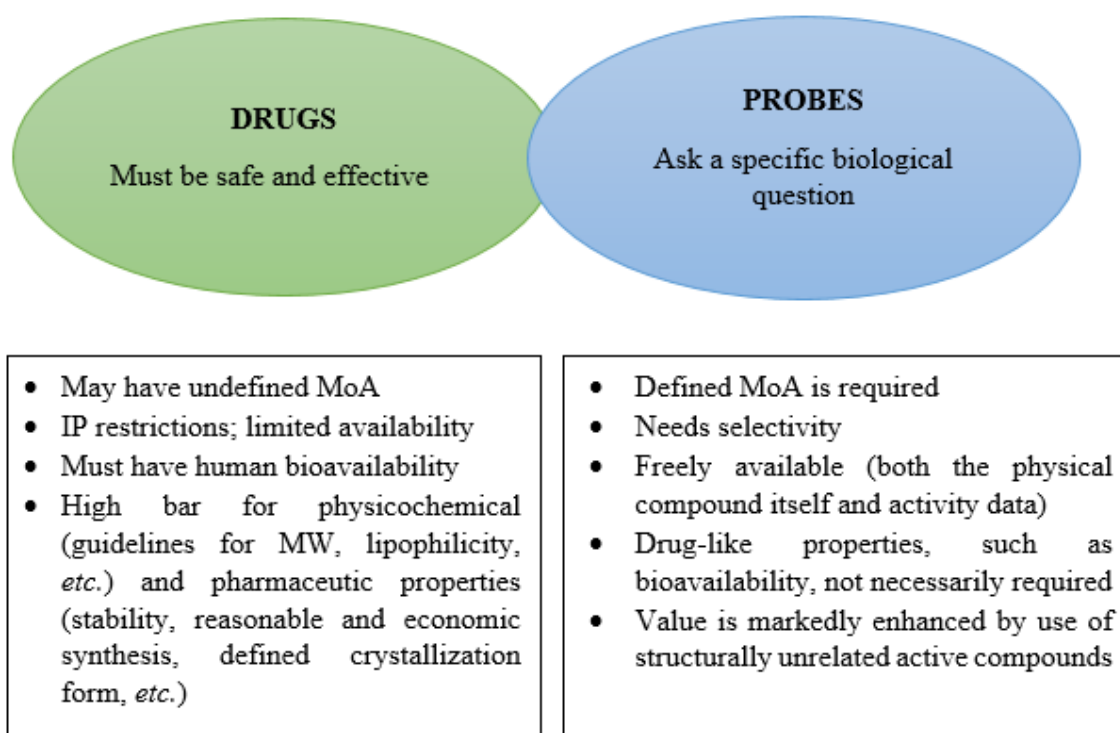


Figure 10; Different purposes for chemical probes and drugs



In addition, chemical probes need to be soluble and stable in water, have good cell permeability, show evidence of target binding and be able to modulate a biological response. Also, it is desirable that the probe is available and pure for researchers.

### **1.6 Fragment-based drug discovery (FBDD)**

The past few decades have witnessed technological advances that can accelerate the pace of drug development.

For many years, HTS campaigns were the most used methodologies in drug discovery for identification of novel hits. HTS involves *in vitro* screening of a large number of structurally diverse compounds (300–500 Da) in a highly automated process. HTS has been successfully applied and has facilitated the development of highly potent and specific inhibitors but, despite of its proved efficiency, suffers from several drawbacks:

- The time and cost associated with the HTS process is high.
- The efficiency tends to decline with increasing size of a library.
- Not able to identify suitable starting points for targeting protein-protein interactions (PPIs).

In contrast to HTS, FBDD relies on obtaining structural information to rationally design or grow small molecules. The relatively low cost of a fragment library and the rapid and efficient identification of hit fragments has created great expectations for this technique. HTS and the development of FBDD will be further explained in Chapter 15 (FragLites).

### **1.7 Biophysical Assays for validation of hits and lead compounds**

Several biochemical and biophysical techniques, including NMR, X-ray crystallography, surface plasmon resonance (SPR), isothermal titration calorimetry (ITC), thermal shift and fluorescence correlation spectroscopy-based assays have been developed to validate the hits obtained by different screening techniques. Each of these assays has advantages and disadvantages, and usually are complementary to each other to afford optimal outcomes. Several of these assays will be used and explained during this thesis.

## Chapter 2: Introduction to allosteric inhibitors of the Ras-activating exchange factor Sos<sup>18</sup>

There is constant communication between cells in our body. The signals they receive and their environment determine when they proliferate, differentiate or die. Our cells communicate with each other using messengers (hormones, growth factors, chemokines) that can deliver information by interacting with other cells. Many of these messengers are too big to cross cell membranes and for this, there is a mechanism called **signal transduction**. This allows membrane receptors to transmit the messages from the surface to the nucleus via proteins inside the cell.

The mutation of these proteins can alter the original message, meaning that the cell nucleus will receive the wrong information. This can be the cause of many diseases, among them cancer. For this reason, the genes that code for signal transduction proteins, activation of which leads to cancer, are known as oncogenes.

Through expression of the Ras proteins, the Ras oncogenes are responsible of delivering signals from the cell surface through receptor tyrosine kinases (RTKs). Ras signaling activates several pathways effecting mitogenic functions (cell division and proliferation). The constitutive activation of these signals through mutation of Ras genes is implicated in many cancer types. Importantly, half of all colon cancers and 90% of pancreatic carcinomas harbor activating Ras mutations. Ras employs a bound guanine nucleotide to create a switch to toggle between “on” and “off” states. In the “off” state, Ras is bound to GDP, but an interaction with a guanine nucleotide exchange factor (GEF) such as Sos (Son of Sevenless) can make Ras change its conformation and liberate the GDP, after that, and because the concentration of GTP in the cell cytosol is much higher than GDP, Ras binds to GTP. This causes a change in the structure of the protein, which initiates the cascade of mitogen-activated protein kinases (MAPKs) (and others) which then can phosphorylate a final target (a transcription factor) delivering the message to the nucleus of the cell. To deactivate itself again, Ras proteins hydrolyze GTP to GDP assisted by GTPase-activating proteins (GAPs), regulating the amount of activated Ras in the cell (Figure 11).

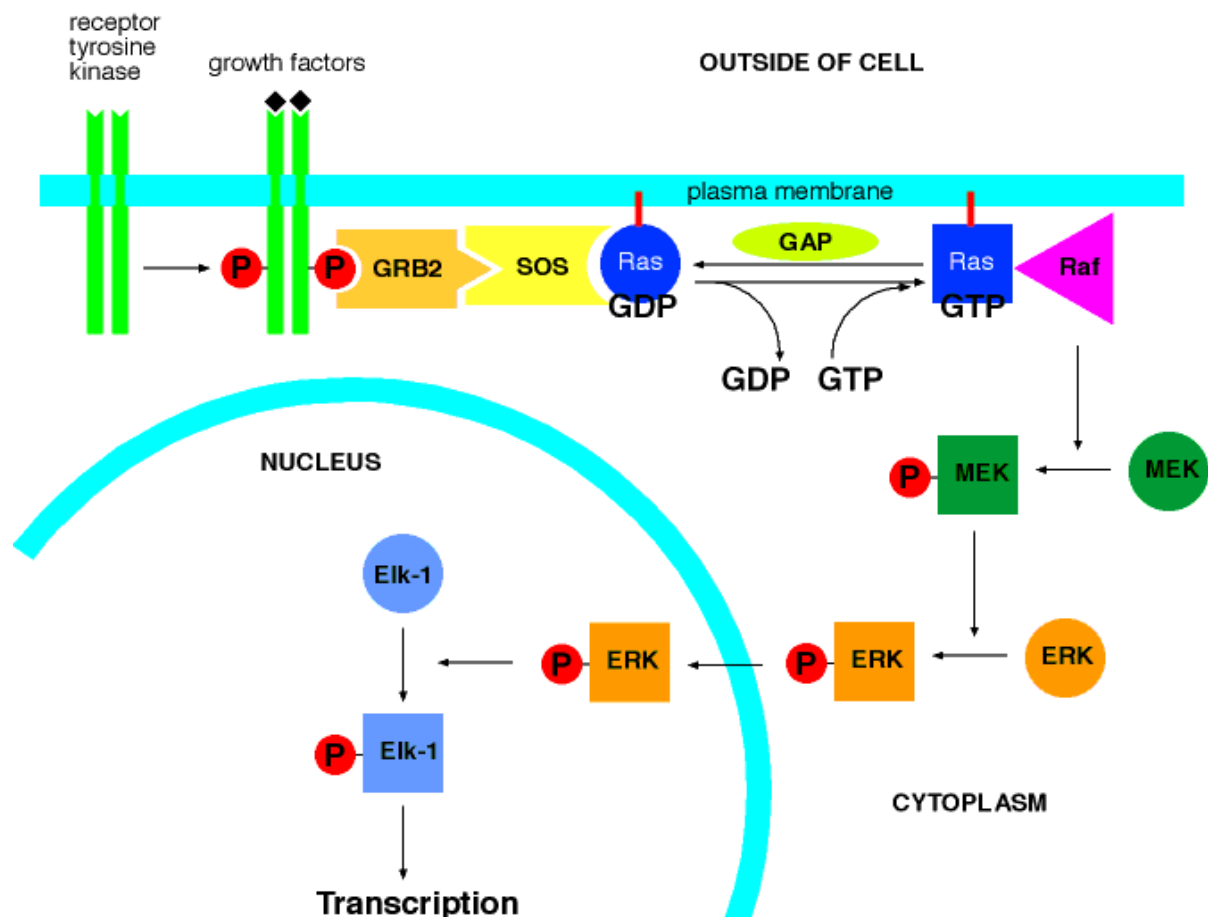


Figure 11; RAS-RAF-MEK-ERK signaling cascade

If Ras suffers certain mutations, the regulation of this mechanism will be disrupted. If Ras is always activated (switched “on”) the message will be delivered continuously, leading to continuous expression of proteins and progression to cancer.

## 2.1 A Ras History<sup>19</sup>

Ras was identified during the studies of acutely transforming retroviruses isolated from mice, rats, cats, monkeys, chickens and turkeys. These viruses have the ability to cause the formation of sarcomas and transform cells in culture with high oncogenic properties.<sup>20</sup>

During the 1960s discoveries of the Harvey murine sarcoma<sup>21</sup> and lately of the Kirsten murine sarcoma viruses<sup>22</sup> provided the basis for establishing the human H-Ras and K-Ras oncogenes respectively. Later in the 1970s Scolnick *et al.*<sup>23,24</sup> performed studies determining the cellular origin of the viral H-Ras and K-Ras genes, and the proteins encoded by those that bind to GDP and to GTP.<sup>25</sup> In these studies was also found that Ras proteins are associated to the inner cell membrane. It was not until the 1980s that a third transforming human Ras gene

was discovered in neuroblastoma-derived DNA and was designated as N-Ras. The three human Ras genes share 82-90% overall sequence identity. Residues 1 – 164 comprise the G domain that contains six conserved sequence motifs shared with other Ras superfamily and GTP-binding proteins. These motifs are involved in binding either phosphate/Mg<sup>2+</sup> or the guanine base of GDP and GTP.<sup>26, 27</sup>

One of the crucial discoveries made by Scolnick was that Ras proteins are preferentially bound to GTP when overexpressed<sup>28</sup> and, as a consequence of the discovery of mutational activated Ras genes in human cancer in 1982, a huge research effort has ensued since then to elucidate the function of Ras protein and the relationship between mutated Ras and cancer. Most of this effort has been focused not only on the complex biochemical mechanisms ruling the action of the Ras proteins, but also on the structure of the proteins and its molecular switch that toggles from “on” to “off” depending on being bound to GTP or GDP respectively. The clear goal since that moment (and still nowadays) is to develop anti-Ras drugs for the treatment of cancer.

The complexity of this target led to an immense amount of work and research, which resulted in several discoveries in the Ras field that can be applied at the same time in other areas of science. Thanks to Ras proteins and their mechanisms, many new, and nowadays fundamental, concepts related to cancer in a molecular level were discovered.<sup>29</sup> Remarkably, the understanding of biochemical mechanisms related to signal transduction was better understood, and this gave insights in how the cell responds to extracellular stimulation. Furthermore, the recently discovered “on-off switches” for different signaling cascades also helped with the elucidation of canonical pathways for signal transduction.<sup>30</sup>

Ras proteins are also the founding members of the large family of *GTPases*, regulators of many of the basic cellular mechanisms.<sup>31</sup> The human family of these versatile regulators could be divided in relation to their sequence identity into different branches: Ras, Rho, Rab, Arf and Ran.<sup>32</sup> The confusing nomenclature of small *GTPases* comes from the tradition of naming them in a way that acknowledges their relationship to Ras (Figure 12).

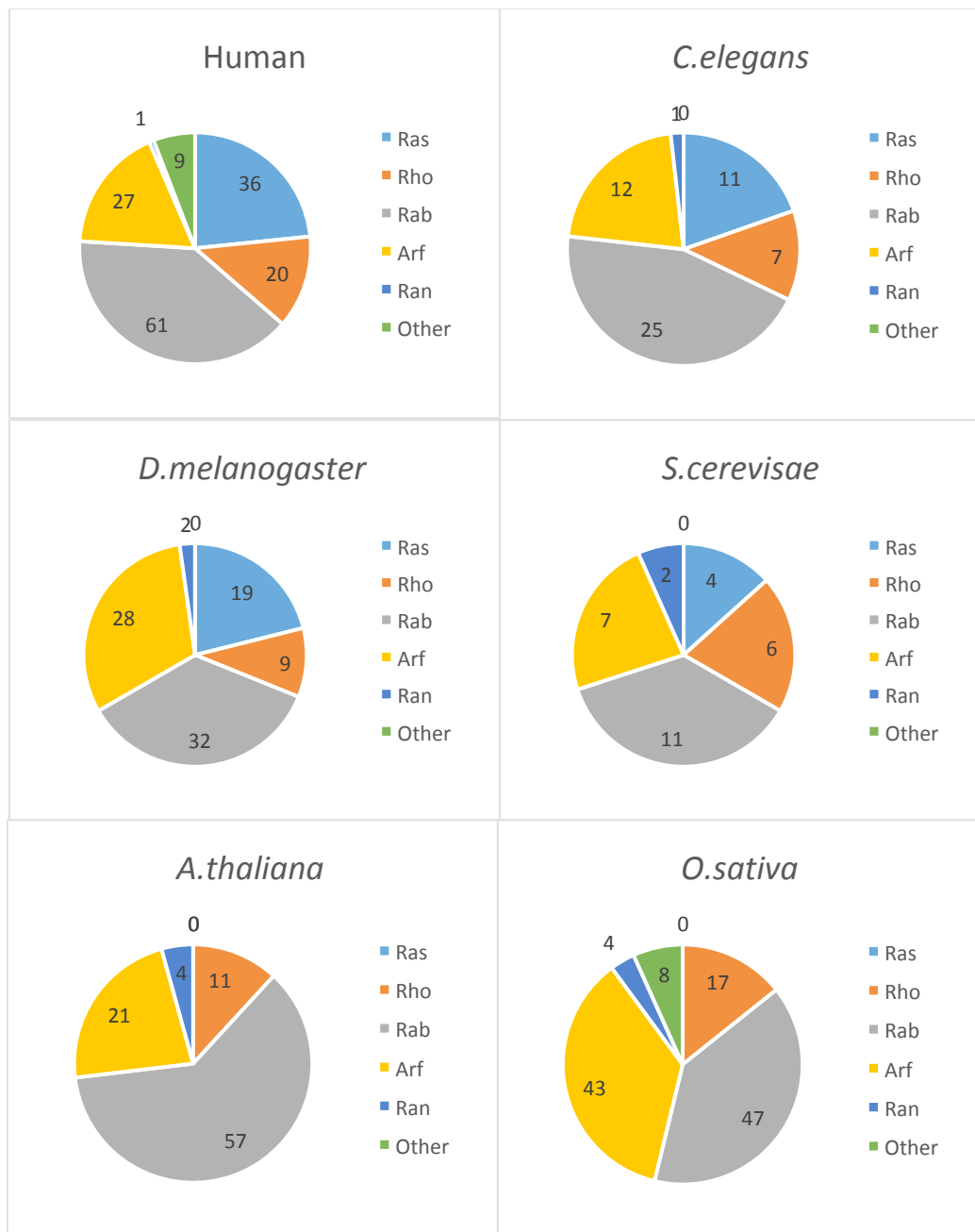


Figure 12; Ras superfamily proteins (numbers indicate total members per family)

More recently, thanks to the sequencing of the human genome, it was possible to identify all the genes encoding Ras-related small GTPases,<sup>14,15</sup> fueling the speculation of other members being involved in cancer disorders. The research showed that many Ras superfamily proteins, despite not being mutated, are critical contributors to oncogenic growth.<sup>33</sup> Ras homologue enriched in brain (Rheb) was found to be an activator of m-TOR and also to be activated downstream of the signaling cascade associated to PI3K.<sup>34</sup> Other members of the Ras superfamily also provided insights into the understanding of normal cellular processes and their regulation.

Regarding the importance of Ras genes in relation to cancer, one of the paramount discoveries was the high frequency of Ras mutations found in colon,<sup>35,36</sup> lung,<sup>37</sup> and pancreatic<sup>38</sup> cancers, three of the four most deadly cancers and with the worst prognosis for patients. Not all the isoforms are likely to be mutated. Mutations are found mainly in K-Ras (85%) and less commonly in N-Ras (12%) and H-Ras (3%). Interestingly, mutations in the different isoforms often happen in the same residues. Missense mutations in H-Ras, K-Ras and N-Ras happen mainly in three residues, G12, G13 and Q61. The most commonly mutated is K-Ras and the most common mutation in this isoform is in the G12 residue (Figure 13).

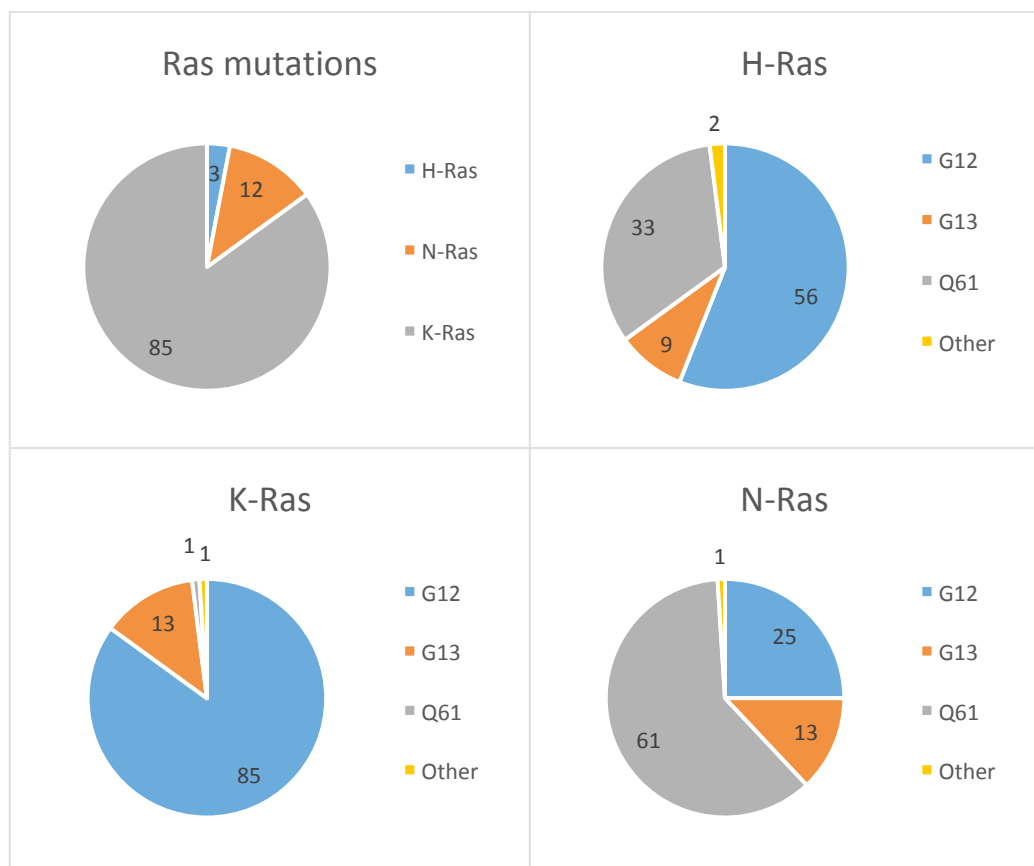


Figure 13; Distribution of Ras missense mutations in cancer<sup>39</sup>

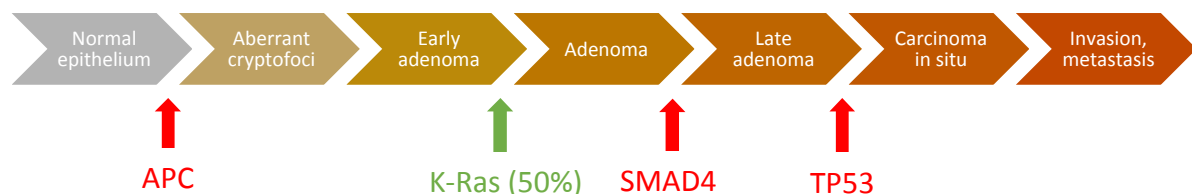
In 1987 it was confirmed by mouse models that Ras mutation can cause cancer but needs cooperating genetic events. This requirement is also true in human colon and pancreatic cancers, and the accumulation of cooperative genetic mutations was also observed.

Thanks to current advances in sequencing technology, the tremendous genetic complexity of the cancer cell genome has been determined. After the sequencing of breast,<sup>40</sup> colon,<sup>41</sup>

pancreatic cancer<sup>42</sup> and glioblastoma<sup>43,44</sup> it was observed that most of cancers happen through a combination of very frequent mutations of specific genes in cooperation with the mutation of a set of less commonly mutated genes specific of each type of cancer. Interestingly, it has also been observed that about 15-20 driver mutations contribute in general to cancer progression.

As we can see in Figure 14, in the case of colorectal cancer, colonic epithelial cells undergo a transition from normal to malignant state. This transformation is driven by specific genetic events, including inactivation of tumor suppressors (**APC**, **SMAD4** and **TP53**) and activation of the **K-Ras oncogene**.<sup>45</sup> Also in Figure 14, in the case of pancreatic ductal adenocarcinoma (PDAC), K-Ras gene mutations are present in 90% of patients. In PDAC cancer, progression of the activating point mutation in the **K-Ras oncogene** occurs early. Inactivation of the p16/**INK4A** gene occurs at an intermediate stage, and inactivation of the **TP53**, **SMAD4**/DPC4 and BRCA2 genes occurs relatively late.

#### Colorectal Cancer:



#### Pancreatic ductal adenocarcinoma (PDAC):

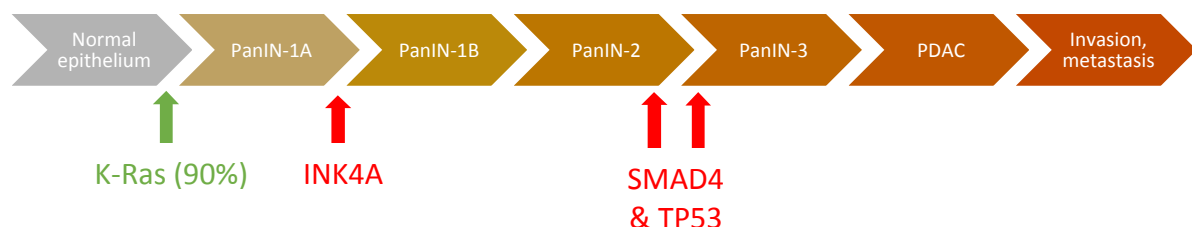


Figure 14; Colorectal and pancreatic cancer progression and gene mutations

The ability of viral Ras proteins to bind GDP or GTP gave the first clues to understand their mechanisms of action and their biochemical roles.<sup>46</sup> This information prompted the idea that Ras proteins may be binary switches regulated by GDP and GTP.<sup>47</sup>

With the finding that the mutant version of the Ras protein is impaired approximately 10 fold in its capacity for hydrolyzing GTP, it was assumed that the aberrant function in cancer cells could be associated with preferential binding to GTP.<sup>48,49,50</sup> Despite the confirmation of this hypothesis, it was again the cooperation with another protein that established the full picture. Trahey and McCormick<sup>51</sup> discovered in 1987 that the GTPase activating proteins (GAPs) activity was responsible for a 300 fold acceleration of the hydrolysis of GTP bound to normal Ras, but not to tumor-associated mutant Ras proteins. Later, in 1990 studies regarding a tumour suppressor protein involved in neurofibromatosis type I,<sup>52,53</sup> showed the existence of a different Ras-GAP, neurofibromin (Nf1).<sup>54</sup> This discovery led to the search and identification of more Ras-GAPs.<sup>55</sup> These findings, together with the established idea that the GTP-bound form of Ras is the activated state, concluded that the key biochemical defect of mutant Ras proteins is GAP insensitivity, resulting in impaired GAP-stimulated GTP hydrolysis that turns into persistent accumulation of the active GTP-bound Ras protein.

The crystal structure of H-Ras provided insights into how Ras functions as a GDP/GTP-regulated switch. Future structural determinations of Ras complexed with GAPs, GEFs and Ras-interacting domains of effectors revealed the complex mechanism of the switch.<sup>56</sup>

Another important step in the understanding of the role of Ras signal transduction was to locate Ras downstream of EGFR and other RTKs located on the cell surface. The first evidence to support it was that EGFR stimulation increased Ras-GTP binding.<sup>57</sup>

Further studies in *C.elegans* and *Drosophila* confirmed the linkage between Ras and EGFR with the discovery of Grb2. Several studies showed that Grb2, through SH2 domains, associates with Sos, completing the connection between EGFR and Ras (Figure 15).



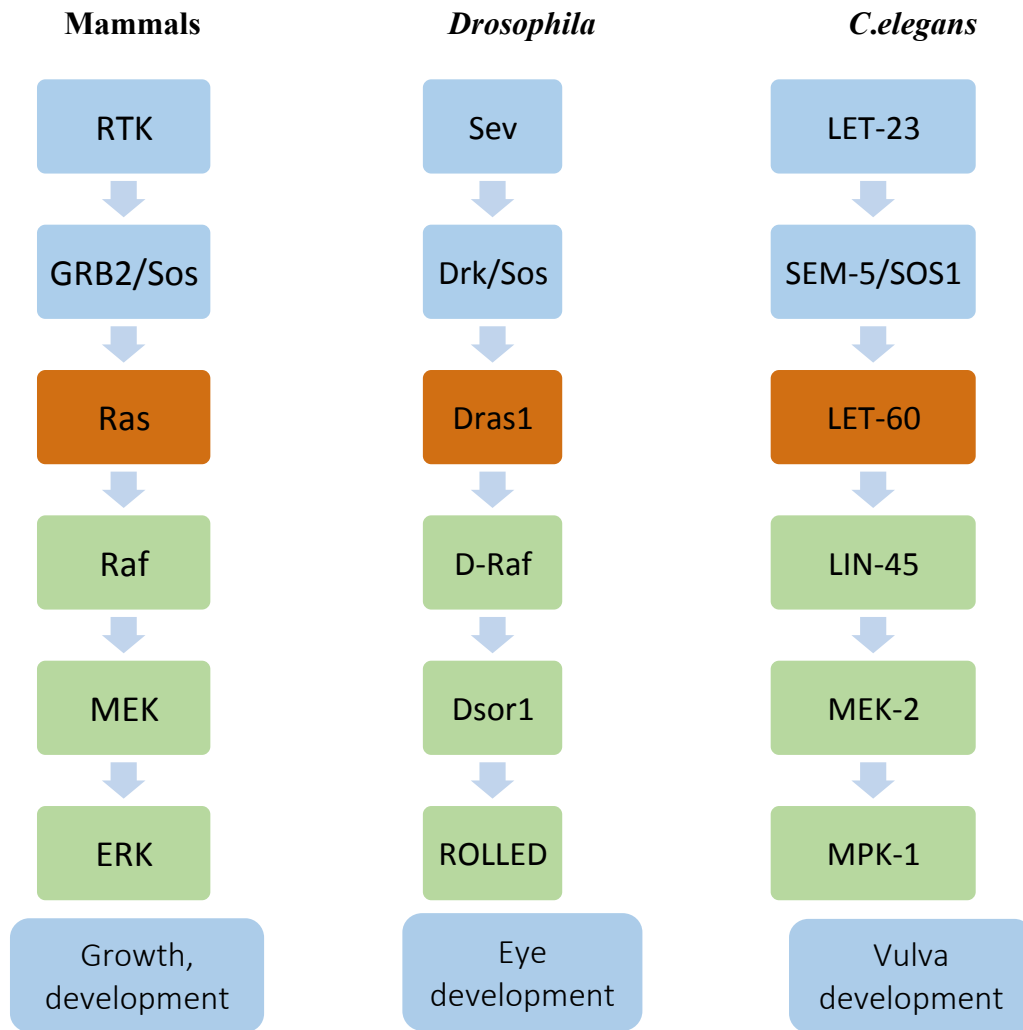


Figure 15; Conservation of Raf-MEK-ERK cascade downstream of Ras<sup>58</sup>

Sos is a guanosine exchange factor (GEF) that induces a conformational change in Ras by forming a complex. This structural change allows deactivated Ras to liberate GDP and associate to GTP to become activated Ras<sup>59, 60</sup>. The complex Grb2:Sos, in association with activated EGFR, promotes Ras activation by the displacement of the cytosolic Sos to the inner membrane. The location of Ras was supported by studies showing that membrane-targeted Sos activates Ras.<sup>61</sup> This observation was fundamental to understand an important concept in signal transduction: “Regulation of subcellular **location** is an important mechanism to regulate the **function** of signaling protein”.

In 1993, after a deep exploration of Ras signaling in *Drosophila*<sup>62</sup> and in *C.elegans*<sup>63</sup> together with information on the biochemistry of Ras proteins determined in mammalian cells, identified Raf serine/threonine kinases as proteins that bind preferentially to activated Ras-GTP, suggesting that this interaction plays a role in selective transmission of signals from the

active form of Ras. This association, together with previous studies supporting Ras- and Raf-dependent activation of the ERK1 and ERK2 Mitogen-activated protein kinases (MAPKs) and the ability of Raf to activate MEK1 and MEK2,<sup>64</sup> defined the Raf-MEK-ERK protein kinase cascade downstream of Ras.<sup>65</sup> The conservation of the Raf-MEK-ERK cascade downstream of Ras in both *Drosophila* and *C.elegans*, suggested that the map of the Ras effector signaling cascade was achieved; however, other Ras effectors were found.<sup>66</sup>

By 1994, the p110 catalytic subunits of class I-PI3K were also recognized as validated Ras effectors<sup>67</sup> (Figure 16).

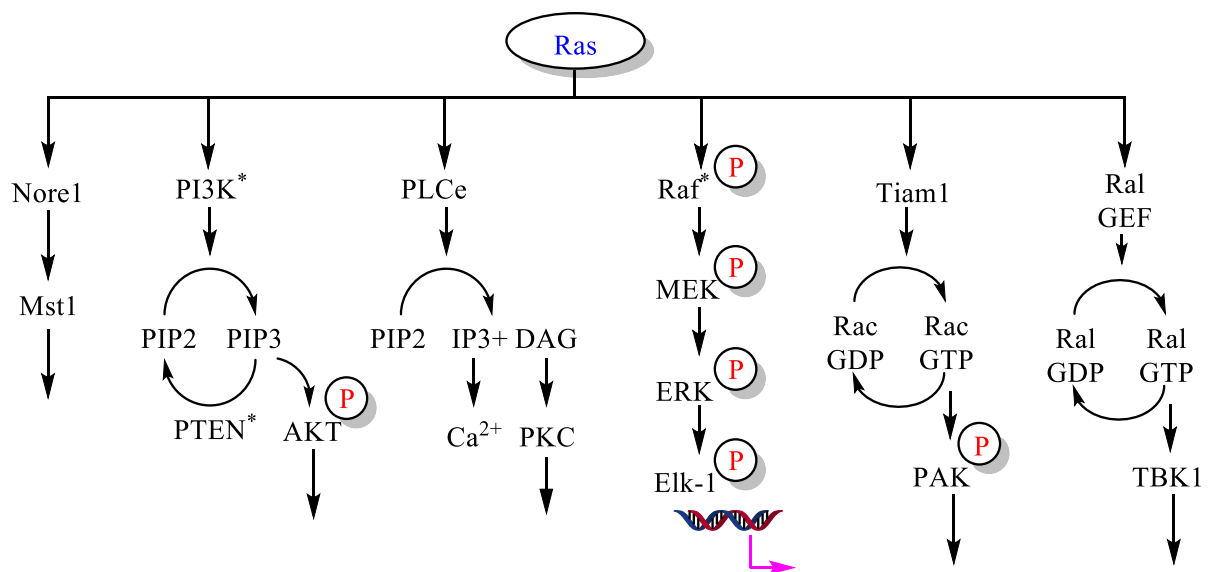


Figure 16; Ras effector pathways implicated in Ras-mediated oncogenesis<sup>68</sup>

A structural property was found to be common to most of the Ras effectors, the presence of a Ras-binding domain (RBD) that promotes association of the effector with GTP-bound Ras.<sup>69,70</sup>

Most of the Ras superfamily proteins have very well-conserved and related isoforms that differ from each other in their “variable” membrane targeting domains rather than in their “constant” G domains. Therefore, Ras activity is not only controlled by its GDP/GTP cycle (concept discovered in 1984<sup>71</sup>), but also by the *C-terminus* of Ras associating with the membrane. Ras proteins must be in the right location to become activated. The importance of the location explains that Ras proteins are synthesized as soluble precursors, which then undergo several steps of post-translational modification, facilitating their transit to the plasma

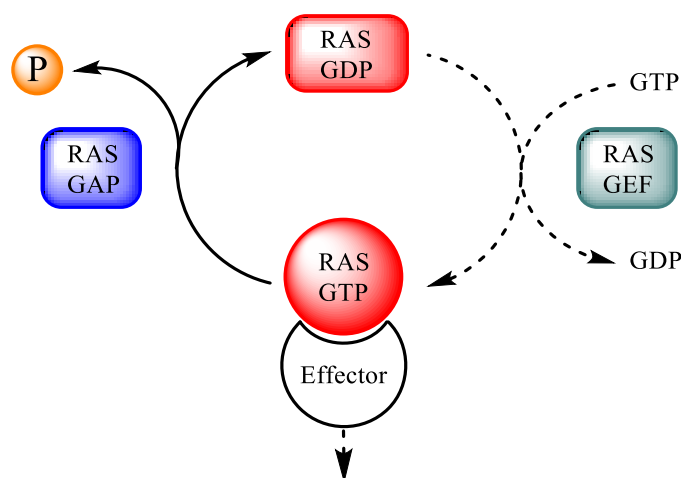
membrane where they are activated by GEFs, inactivated by GAPs and interact with effectors to transmit their signals downstream.

In 1980, an important study using electron microscopy confirmed the presence of H-Ras at the inner leaflet of the cell membrane.<sup>72</sup> In 1989 it was determined that Ras proteins are modified by farnesylation,<sup>73,74</sup> and that this modification is an obligatory first step to allow palmitoylation. Without farnesylation at a very specific residue (Cys 186), Ras proteins become cytosolic. In contrast, Ras proteins that are farnesylated but do not undergo subsequent processing steps are located to endomembranes. The location of Ras proteins when not processed was not clear and it was not until 1999 when Philips *et al.*<sup>75</sup> showed that H-Ras and N-Ras trafficked through the Golgi apparatus on the way to the membrane.

Originally there was a general acceptance that all the Ras isoforms were identical in biological function and the study of H-Ras dominated the scene. This bias towards H-Ras was due to the availability of H-Ras expression vectors, mutant proteins and easy crystallization. But in recent years it has been proved that the most common mutation in human cancers is present in the K-Ras and N-Ras genes, shifting the research towards K-Ras in particular.

## 2.2 Ras Functions<sup>76</sup>

The main function of Ras is to convert GTP to GDP, functioning as a GTPase. Ras functions as a “binary molecular switch”, toggling between “ON” and “OFF” states, based on whether it is bound to GTP or GDP respectively (Scheme 1).<sup>77</sup>



Scheme 1; Ras GTP/GDP cycle regulation by GEFs and GAPs

The function of the switch involves the exchange of GDP for GTP by GEFs, turning the switch to the “ON” state. This should be reversible and to turn the switch to the “OFF” state GTP is hydrolysed by Ras to GDP. This process is very slow and is accelerated by GTPase activating proteins (GAPs).

### 2.3 The structure of the G Domain

The G domain of the Ras protein has the fundamental function of binding the nucleotide and hydrolysing GTP to GDP. This domain has a universal and well conserved structure for all the GNBPs and a universal switch mechanism, which consists of six-stranded  $\beta$  sheets and five helices located on both sides. The most important contributions to binding are made by the interactions of the nucleotide base with the N/TKXD motif (where X is any amino acid) and of the  $\beta,\gamma$ -phosphates with the conserved P-loop (phosphate-binding loop).<sup>78</sup> Specificity for guanine is due to an Asp side chain (from the NKxD motif), which forms a bifurcated H bond with the base, and to a main chain interaction of an invariant Ala (146) with the guanine oxygen (Figure 17).

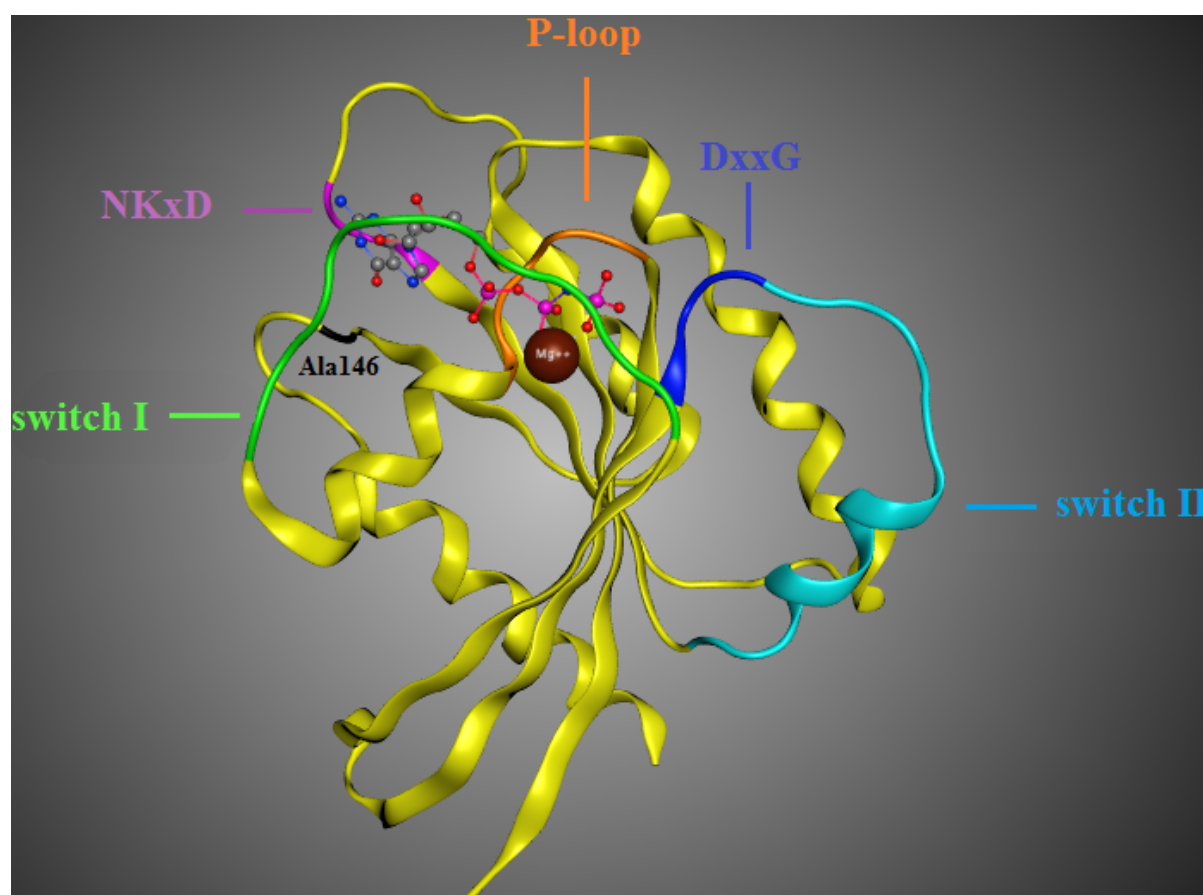


Figure 17; Ribbon plot of the G domain generated using MoE (PDB ID: 3GFT)

Structures of the subunit of heterotrimeric G proteins like  $G_t$ ,  $G_i$ , and  $G_s$ , and several Ras-related proteins like Rap, Ral, Rac, Rho, Cdc42, Arf, Arl, Rab and Ran have been also determined.<sup>79, 80, 81</sup>

### 2.3.1 The Structural Changes of the switch

Whereas the GDP-bound proteins show variations in their structures, the GTP bound forms of the G domain are remarkably similar.

The trigger for the conformational change is universal. In the GTP-bound form, there are two hydrogen bonds from  $\gamma$ -phosphate oxygens to the main chain NH groups of the invariant Thr35 and Gly60 residues in switch I and II, respectively. The conformational change could be described as a loaded-spring mechanism where release of the  $\gamma$ -phosphate after GTP hydrolysis allows the two switch regions to relax into the GDP-specific conformation (Figure 18).

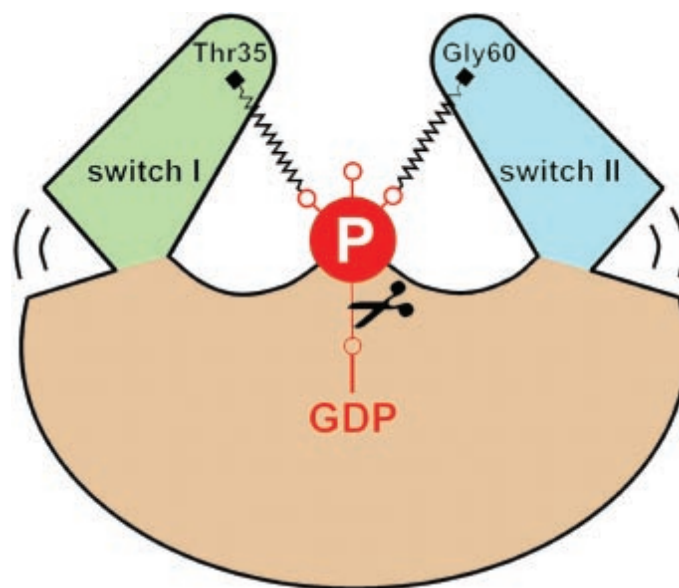


Figure 18; The universal switch mechanism<sup>76</sup> (Reprinted with permission from AAAS)

### 2.3.2 Moving out the nucleotide, GEF reaction

GEFs help by accelerating the release of GDP from Ras, which is slow by nature. This is a multistep process involving several fast reactions. This action could be summarized as a binary protein-nucleotide complex that leads to another binary nucleotide-free complex via a ternary GNBPs-nucleotide-GEF intermediate. These reactions come to an end by binding

GTP, which is predominant in the cell. The fact that the cycle can start again makes GEFs act as catalysts to unbalance a specific equilibrium<sup>82</sup> (Scheme 1).

The complexes formed between GEFs and their GNBPs have several features in common, suggesting mechanistic similarities. The GEFs can interact with both switch I and II regions, and more importantly, insert residues close to the P loop and the  $Mg^{2+}$ -binding area. This results in a conformational change that inhibits the binding of phosphates and  $Mg^{2+}$ .  $Mg^{2+}$  in this case plays an important role, however studies showed that the  $\beta$ -phosphate-P loop interaction is the most important interaction for binding of the nucleotide. The structural change in the P loop is probably the reason for the decreased affinity towards GTP.  $Mg^{2+}$  is pushed out of its position by elements of switch II itself (see Ala59) or from residues of the GEF. The residues of the P-loop suffer a conformational change and the lysine associated to those is oriented now to the carboxylates of Glu62. Switch I is pushed out, and switch II pulls toward the nucleotide-binding site (Figure 19).

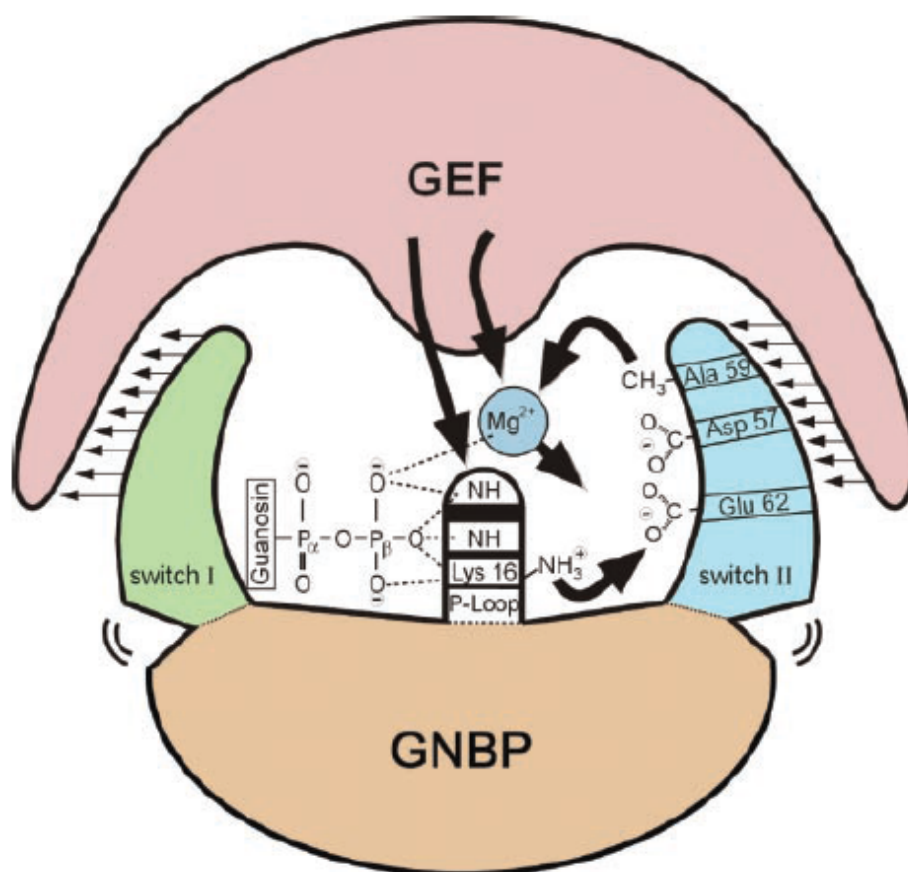


Figure 19; GEF:GNBP interaction<sup>76</sup> (Reprinted with permission from AAAS)

### 2.3.3 Son of Sevenless (Sos)<sup>83</sup>

Sos is a major member of the GEF family<sup>84</sup> and it is a multidomain enzyme with a Ras Exchanger Motif (Rem) domain and a Cdc25 domain that are both essential for Ras GTPase activation and signalling (Figure 20).

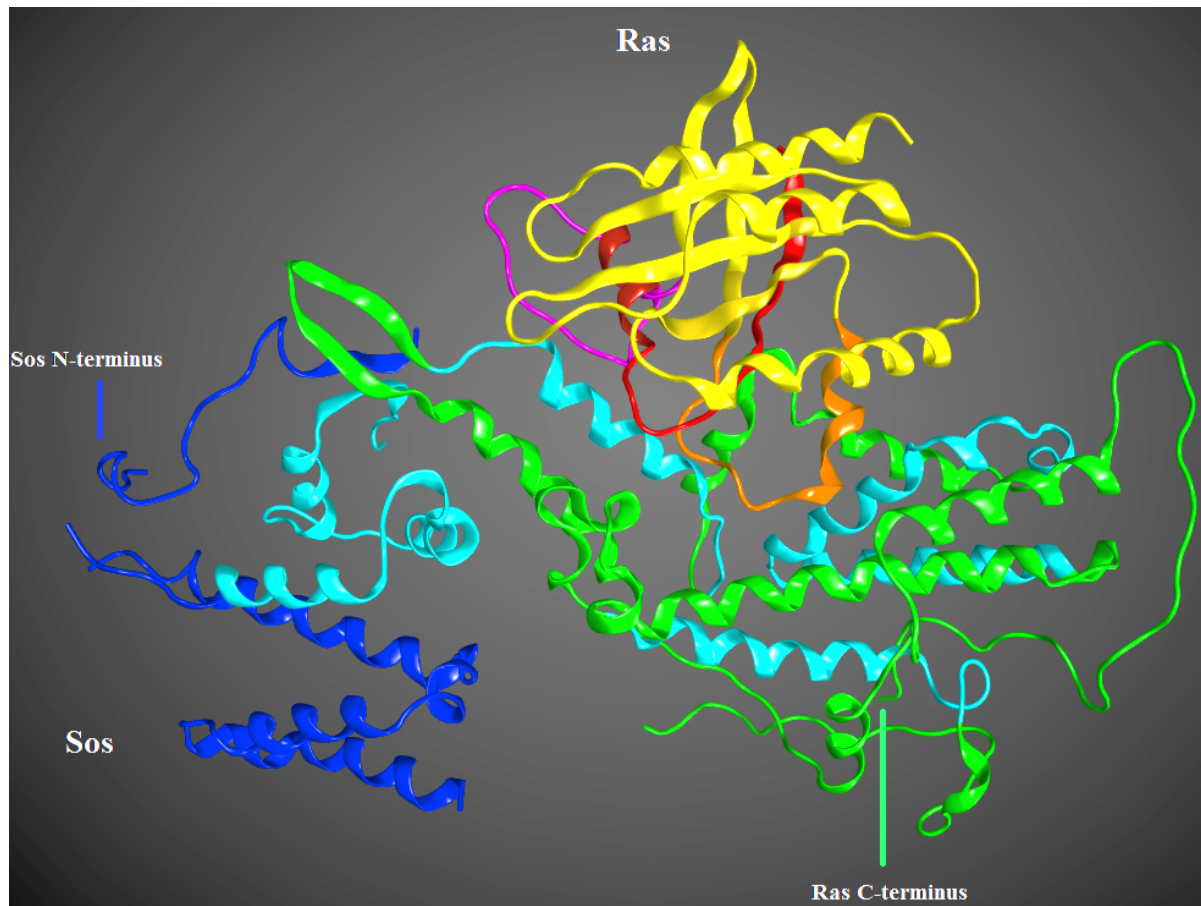


Figure 20; Ribbon plot of SoS:Ras complex generated using MoE (PDB ID: 1BKD); Sos N-domain (blue), Sos catalytic domain (green); Ras Switch I and II (orange and magenta), P-loop (red), conserved regions among Ras family exchange factors (cyan)

Cdc25 is the catalytic site, where Ras-GDP binds. The Rem domain is an allosteric site. Binding of Ras-GTP to the distal allosteric site results in a ternary complex Ras(GDP):Sos:Ras(GTP) which increases Ras exchange activity.<sup>85</sup>

The formation of the Ras-Sos complex causes drastic changes in Switch I, which is removed from the binding site, breaking a network of direct and water-mediated interactions between Switch I and the nucleotide, as well as impeding binding of the  $Mg^{2+}$  ion and phosphate.<sup>86</sup> The backbone of Switch II is compressed, helping to occlude the  $Mg^{2+}$  ion and interrupting



the phosphate binding side chains (Figure 21).

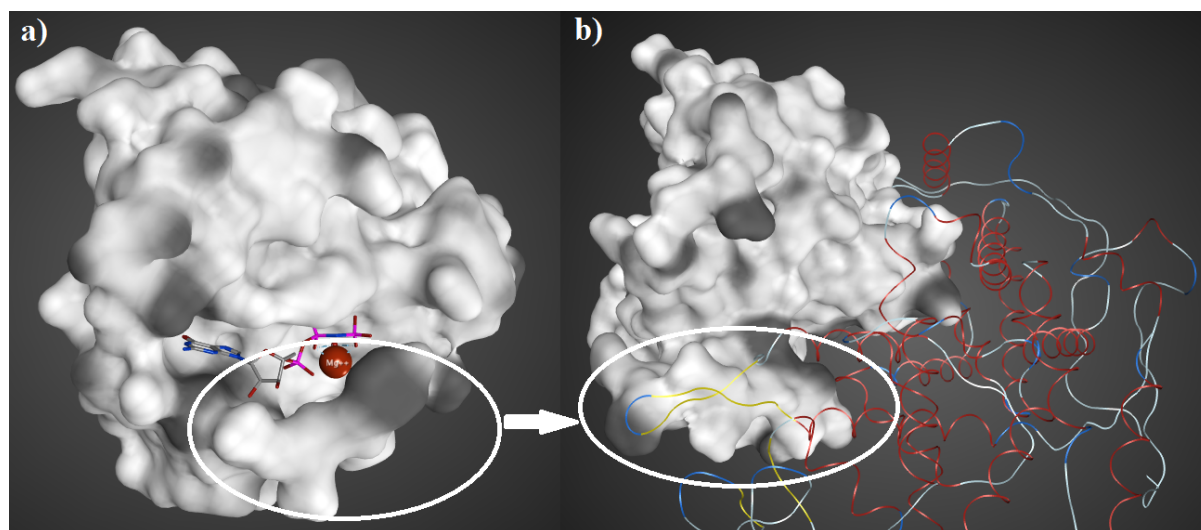


Figure 21; Change in Switch I causing liberation of GDP after interaction with Sos is highlighted (Generated using MoE); **a)** Ras (white ribbon) - GDP analogue (stick representation) interactions (PDB ID: 5P21) to **b)** Ras (white ribbon) -Sos interactions (PDB ID: 1BKD)

Sos mediates nucleotide dissociation from the ternary complex and the formation of the new complex, Ras-Sos is disrupted when GDP or GTP binds to Ras. GTP and GDP bind with similar affinity, but intracellular GTP concentration is over 10 times that of GDP, resulting in an innate preference for the GTP bound form.

## 2.4 Effectors

The effectors of the GTP-binding proteins will be molecules that can interact (or interact more tightly) to GTP bound proteins than to GDP bound proteins. This implies that the effector binding mode will involve the switch regions of G proteins as these regions are the ones in contact with the different nucleotides.

The Ras-binding domain (RBD) of the Raf protein kinase (Ras effector) is a very small domain with a ubiquitin fold, which binds to Ras by forming a GTP sensitive interprotein  $\beta$  sheet between the two molecules.<sup>87</sup> This binding domain is contained in many other proteins, most of them with similar fold and function.

## 2.5 GAP proteins and GTPase Reaction

The GTPase reaction is found to be very slow for all GNBPs, and for most biological processes is required to be complete within minutes after GTP is loaded. For this, GAP



proteins are necessary, and work in assistance of Ras forming a complex. The speed of hydrolysis changes dramatically:

- GTP hydrolysis in water takes 200 days
- GTP hydrolysis by Ras takes 30 min
- GAP assisted GTP hydrolysis by Ras takes 50 ms<sup>88</sup>

Structures of RasGAP and RhoGAP in complex with their respective GNBPs in the presence of  $\text{AlF}_x$  ( $\text{AlF}_3$  or  $\text{AlF}_4^-$ ), show that GAPs supply a so-called arginine finger into the active site.<sup>89</sup> There are some common principles in the mechanism of GAP action on Ras, Rho, and G proteins (Figure 22).

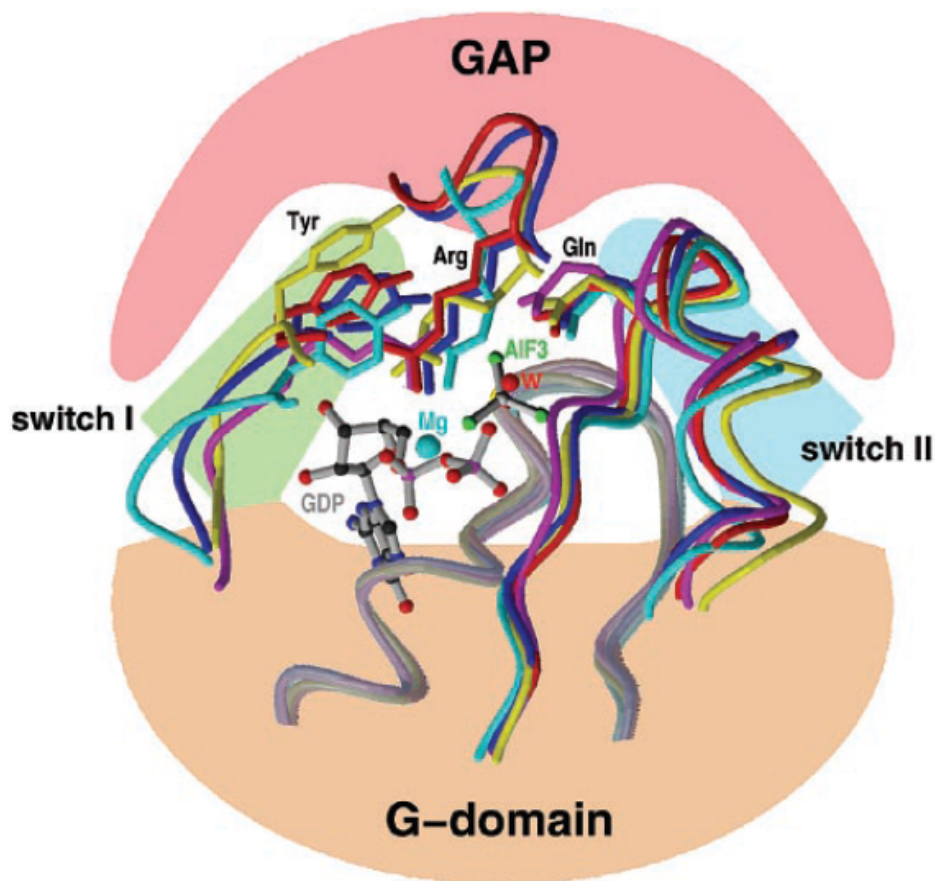


Figure 22; “GAP:G-domain” interaction. This figure was prepared with Molscript.<sup>90</sup> Ras- RasGAP (yellow), Rho- RhoGAP (red), Rac- ExoS (cyan), and  $G_i$ -RGS (magenta); catalytic water (W). All in the presence of GDP and aluminium fluoride compounds ( $\text{AlF}_4^-$  or  $\text{AlF}_3$ )

## 2.6 Is Ras Druggable?<sup>91</sup>

Although the direct targeting of Ras is a very challenging goal due to its rounded surface and lack of potential druggable sites, some potential pockets on the surface of K-Ras have been found, allowing the use of computational approaches<sup>92, 93</sup> to find small molecules inhibitors.

### 2.6.1 Targeting Ras directly

#### Low-affinity inhibitors

Initially the compounds were designed to compete with GTP for the binding site on Ras,<sup>94</sup> some compounds were found to bind close to the region of switch II, however all of these initial series had a hydroxylamine group, which is highly toxic and has poor metabolic properties due to instability. On top of that the potency of those is very low.

Another type of compounds was found to inhibit the formation of the dimeric complex Ras:Raf. These compounds are based in the non-steroidal anti-inflammatory drug sulindac sulphide. The most potent of those compounds (**IND12**), is a selective inhibitor for the proliferation of Ras-transformed-cells, but lacks potency and has many side effects (Figure 23).<sup>95</sup>

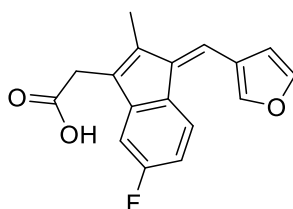


Figure 23; **IND 12**

In a yeast two-hybrid screen for compounds that could potentially inhibit the Ras-Raf interaction, small molecules (like **MCP1**-Figure 24) were discovered that inhibited Raf activation in cancer cells.<sup>96</sup> Again, the compounds are not potent enough. Structural information on how they bind has not yet been obtained but could be a starting point for improving its potency and develop suitable inhibitors for the Ras-Raf interaction.

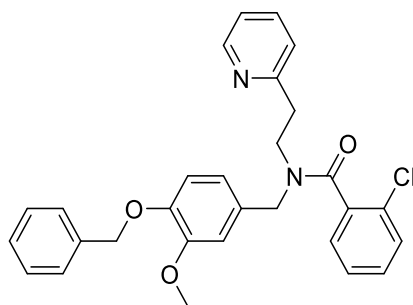


Figure 24; **MCP1** was the product of a yeast screening

$\text{Zn}^{2+}$  cyclen (Figure 25) was shown to bind to Ras and block the interaction with Raf,<sup>97</sup> but this molecule has no selectivity for Ras due to its chelating nature. The side effects limit any further development.

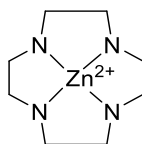


Figure 25;  $\text{Zn}^{2+}$  cyclen

Despite of all these problems, these low-affinity inhibitors revealed the existence of pockets in Ras that could be used to disrupt the interaction with its main effector, Raf, with a small molecule and therefore inhibit the signaling cascade.

### GEF inhibitors

The Ras GTP/GDP cycle is regulated negatively by GAPs and positively by GEFs (Scheme 1, page 39). The most studied and prominent Ras GEF is Sos, which forms a complex with 2 Ras proteins ( $\text{Ras}^{\text{cat}}\text{Sos}:\text{Ras}$ ) using 2 binding sites (allosteric and catalytic).<sup>98</sup>

A peptide based on the  $\alpha$ -H helix of Sos, the region that can bind to Ras (**HBS3 peptide**-Figure 26) was found to bind very closely to the Switch I and Switch II regions. After this discovery, the research for a new peptide with good pharmacokinetic properties started. The resulting pseudo-peptide bound to GDP-bound Ras and inhibited the Ras activation which therefore regulated the ERK-MAPK signaling, however the peptide was again not potent enough and further optimization is needed to lead to drug-like molecules.

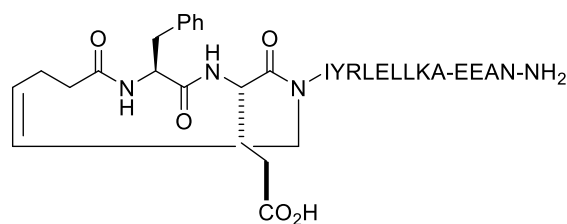


Figure 26; The **HSB3 peptide** is a mimic of the Sos  $\alpha$ -H helix that interacts with H-Ras<sup>99</sup>

In 2012 two compounds were reported to bind to K-Ras after fragment based screenings by two different groups. Genentech could identify a compound named as **DCAI**, which binds to a pocket located between the  $\alpha 2$  helix and  $\beta$  sheets of K-Ras (Figure 27).

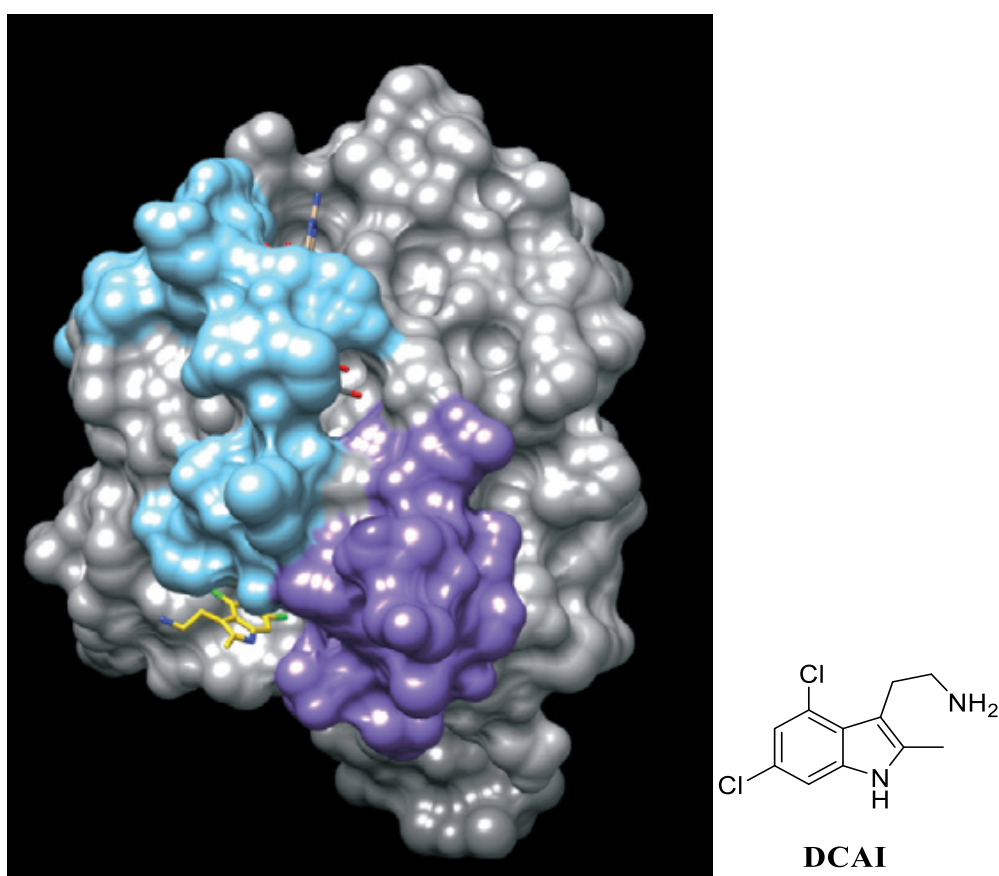


Figure 27; K-Ras<sup>G12D</sup> GDP bound to **DCAI** (PDB ID: 4DST)<sup>100</sup>

**DCAI** inhibited Sos-mediated nucleotide exchange by blocking the interaction between Ras and Sos, and inhibited Ras activation in cells, but again very weakly ( $IC_{50} = 340 \mu M$ ). At the same time another group in Vanderbilt, discovered several compounds using NMR-based fragment screen. The most potent was named **VU0460009**,<sup>101</sup> which binds to the same cleft as **DCAI** (Figure 28).

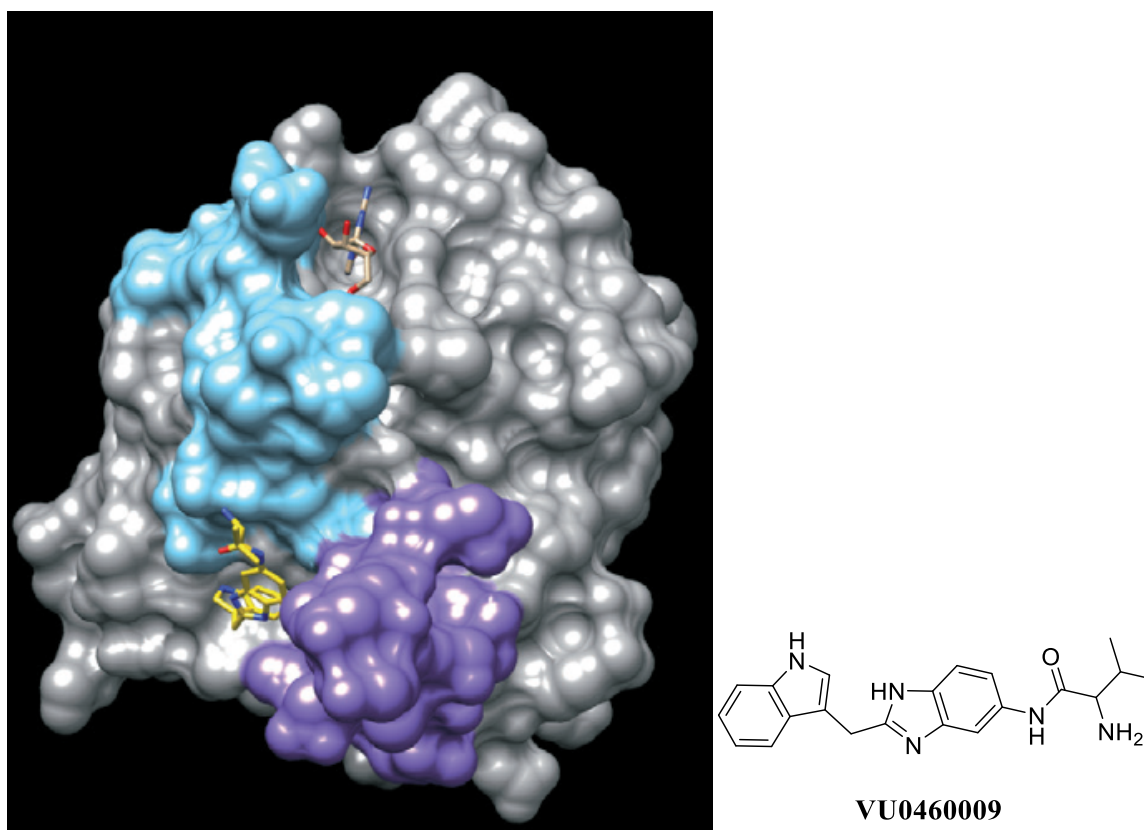


Figure 28; X-Ray structure of K-Ras<sup>G12V</sup>-GDP bound to **VU0460009** (PDB ID: 4EPY)

This pocket cannot be observed in the ligand free form of K-Ras, it is only with the conformational change induced by binding of the compounds that this pocket becomes available. This new conformation inhibited the Sos-mediated nucleotide exchange by blocking the Ras-Sos complex formation but like many other compounds, these new small molecules weakly bind K-Ras. The improvement of these compounds is very challenging and it is not sure if the mechanism of action would be valid with mutated Ras proteins.

### Recent publication for the inhibition of the GEF (Sos)<sup>102</sup>

The identification of a potent small molecule inhibitor with the capacity to disrupt the interaction between K-Ras and Sos was undertaken via a dual approach:

- 1) Fragment screen to identify inhibitors via K-Ras-Sos complex stabilization.
- 2) A High-throughput Screen (HTS) designed to search for inhibitors of the enzymatic Sos nucleotide exchange activity, via binding either to K-Ras or to Sos.

For the first approach the aim was to identify molecules able to stabilize the interaction between K-Ras and Sos and drive the complex to a dead-end. An NMR fragment screen for

binders of the complex was performed to identify potential stabilizers (Figure 29).

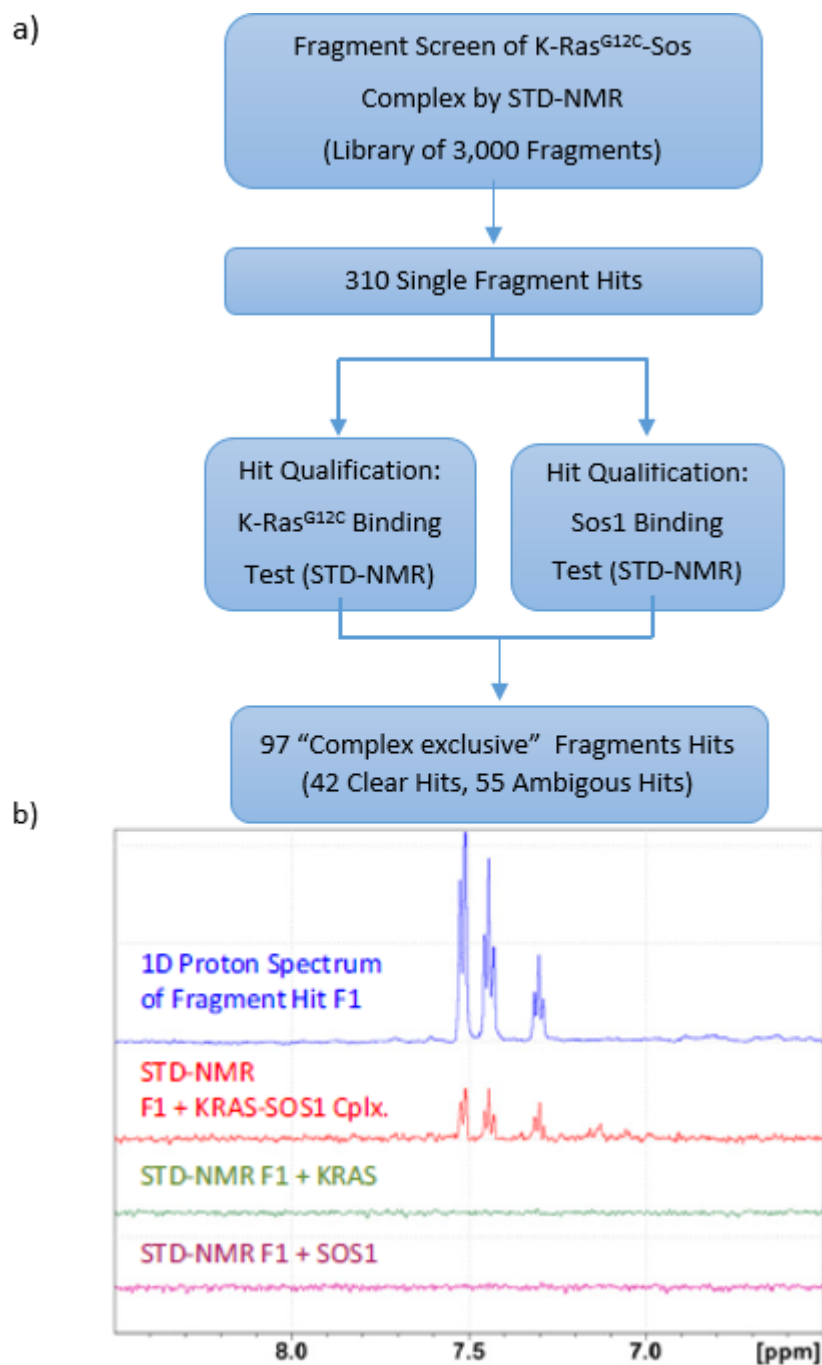


Figure 29; a) Screening cascade; b) example spectra for (8) 1D proton spectrum (blue) and STD-NMR spectrum with K-Ras<sup>G12C</sup>-Sos1 complex (red), K-Ras<sup>G12C\_C118S</sup> (green), and Sos1<sup>cat</sup> (purple)<sup>102</sup>

A library of 3000 fragments was screened by STD-NMR yielding 310 single fragment hits. The hits found were subjected to K-Ras and Sos screens. Of the 97 “complex exclusive” fragments (fragments binding exclusively to the K-Ras<sup>G12C</sup>:Sos<sup>cat</sup> complex) 42 were unambiguous and selected for crystallization. Soaking of the 42 fragments gave 13 co-crystal

structures that showed binding to a mainly hydrophobic pocket on Sos located adjacent to K-Ras.

Of the different fragments identified, compound **(8)** stabilized K-Ras<sup>G12C</sup>-Sos<sup>cat</sup> complex in three different assays: In the 2D NMR assay, this was indicated by the reduction of signals for <sup>15</sup>N-labeled K-Ras<sup>G12C</sup> upon addition of **(8)**. In the SPR assay, addition of **(8)** increased the amount of K-Ras<sup>G12C</sup> binding to immobilized Sos<sup>cat</sup>. In the interaction assay, **(8)** addition resulted in an increased homogeneous time-resolved fluorescence (HTRF) signal. **(8)** interacts with Sos via a  $\pi$ - $\pi$  interaction with Phe890 and forms two hydrogen bonds to Tyr884 and Asp887; moreover, the aminomethyl group forms a cation- $\pi$  interaction with Tyr884 (Figure 30).

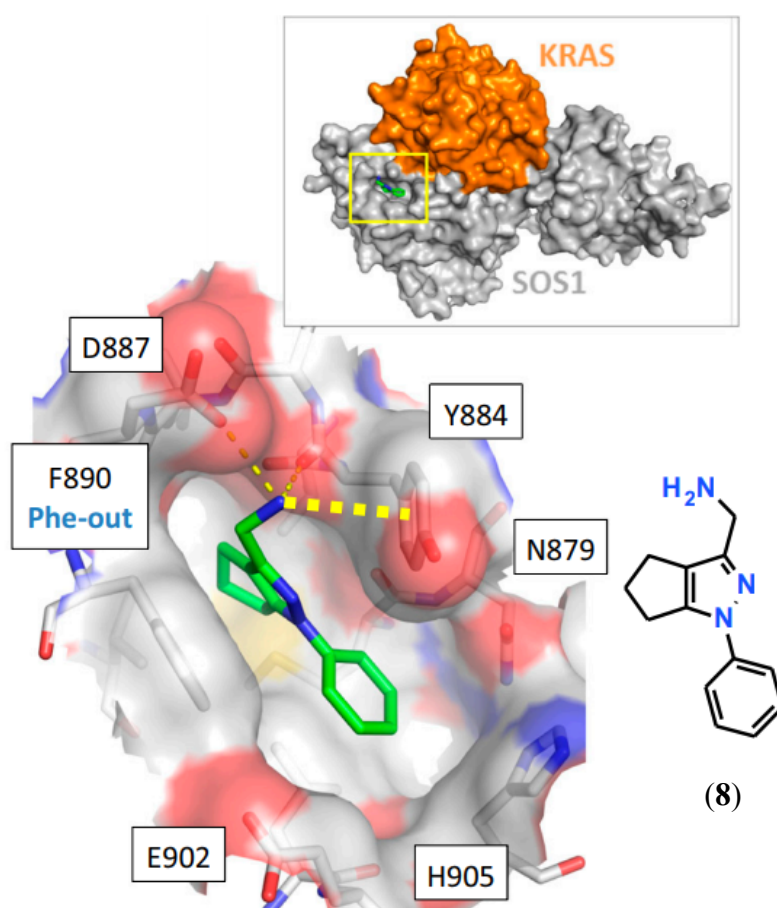


Figure 30; Co-crystal structure of **(8)** bound to K-Ras<sup>G12C</sup>-Sos<sup>cat</sup>. Overall complex with location of the fragment binding site (yellow box); Inset is the area in the yellow box enlarged, showing hydrogen bonds as thin dashed lines and cation- $\pi$  interaction as a thick dashed line<sup>102</sup>

Attempts of SAR exploration on **(8)** were conducted but no significant improvements were observed in terms of potency and the approach was abandoned.



For the second approach, an HTS with a developed enzymatic assay was undertaken with a Bayer library of over 3 million compounds (Figure 31) identifying compound (**9**).

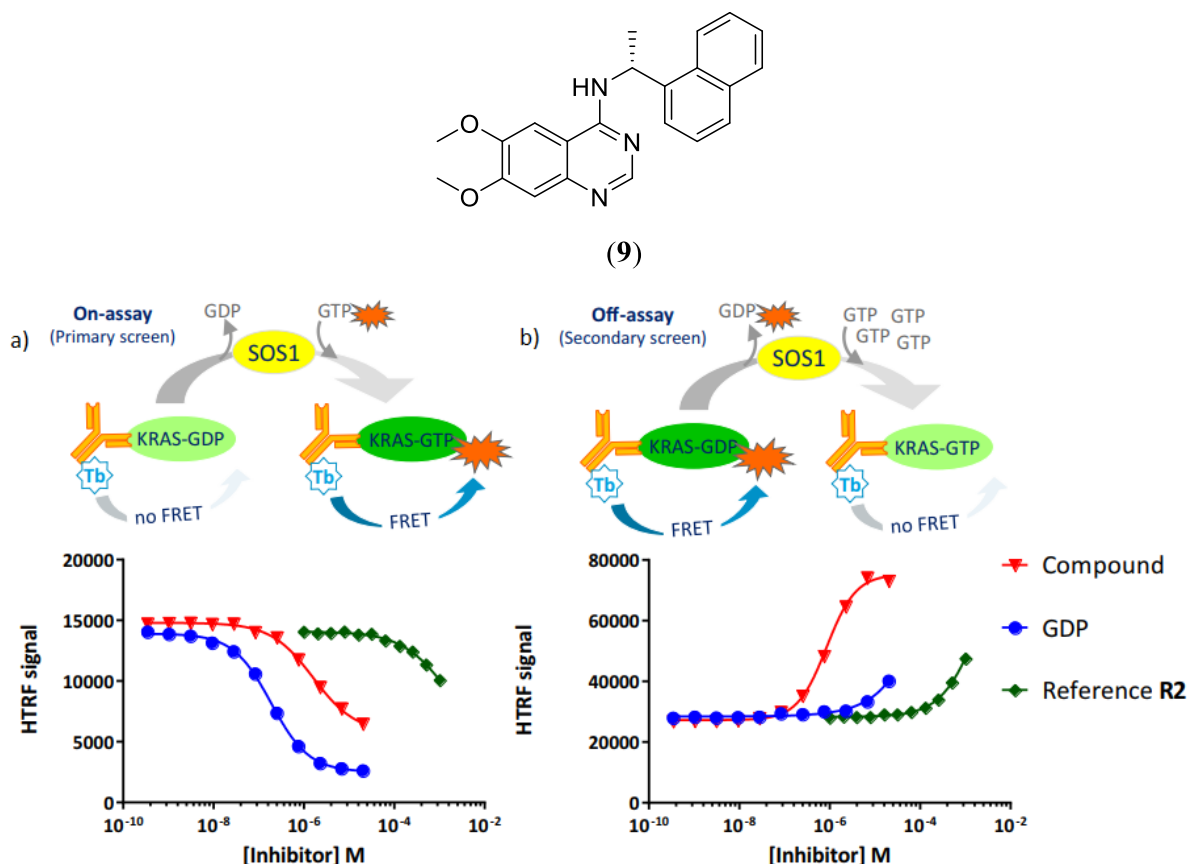


Figure 31; a) Quantifying the Sos-mediated loading of a fluorescently labeled GTP analogue onto K-Ras<sup>G12C</sup>, which results in an increased HTRF signal; b) Monitoring the SOS1-catalyzed HTRF signal decrease by the deloading of a fluorescently tagged GDP analogue preloaded onto K-Ras<sup>G12C</sup>. (K-Ras reference compound (**R2**) is a direct K-Ras inhibitor)

A thermal shift assay was also used to confirm initial hits and to analyse the interaction of the small molecules with K-Ras or Sos. Compound (**9**) stabilized Sos, but not K-Ras (wild type or mutant). ITC also confirmed a strong enthalpy driven binding (-14.1 kcal/mol). Further SAR exploration of compound (**9**) yielded compound (**10**), with an IC<sub>50</sub> of 21 nM. Co-crystallization with (**10**) revealed, unambiguously, that only the (R)-enantiomer (**10**) had bound in the crystal (Figure 32).



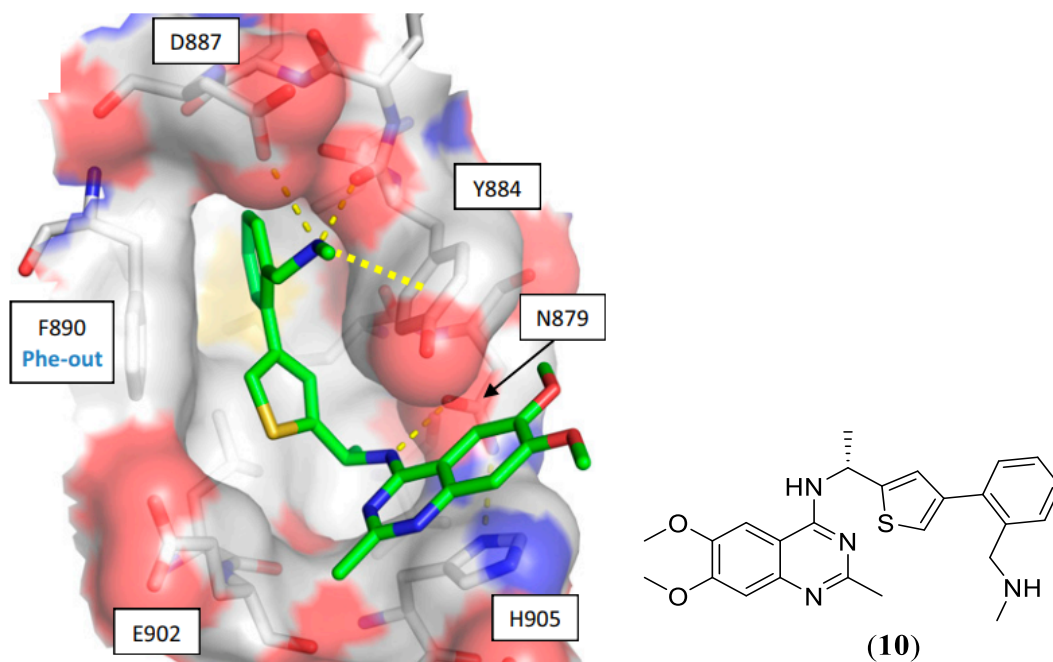


Figure 32; Hydrogen bonds shown as thin dashed lines, cation- $\pi$  interaction as a thick dashed line

Thanks to this work a cell-active and selective inhibitor of Sos with a 21 nM potency was reported for the first time. Compound (10) could be used for future investigations of the Ras-Sos interaction *in vitro*. Further improvement of the pharmacokinetic profile of the ligand are being undertaken for viable *in vivo* experiments.

### Inhibiting Raf-1 binding

One compound was found to have the ability to inhibit the Ras-Raf interaction by binding to the Ras binding domain (RBD) known as Raf-1<sup>103</sup> after a computer docking screen of a virtual library, and then a second compound was also found after a similarity search. These compounds named **Kobe** derivatives decreased the phosphorylation of downstream proteins in the Ras pathway and more importantly inhibited the allosteric Ras binding site of Sos. Unfortunately, these molecules contain a thiosemicarbazide, which should be replaced in future developments due to potential toxicity issues (Figure 33).

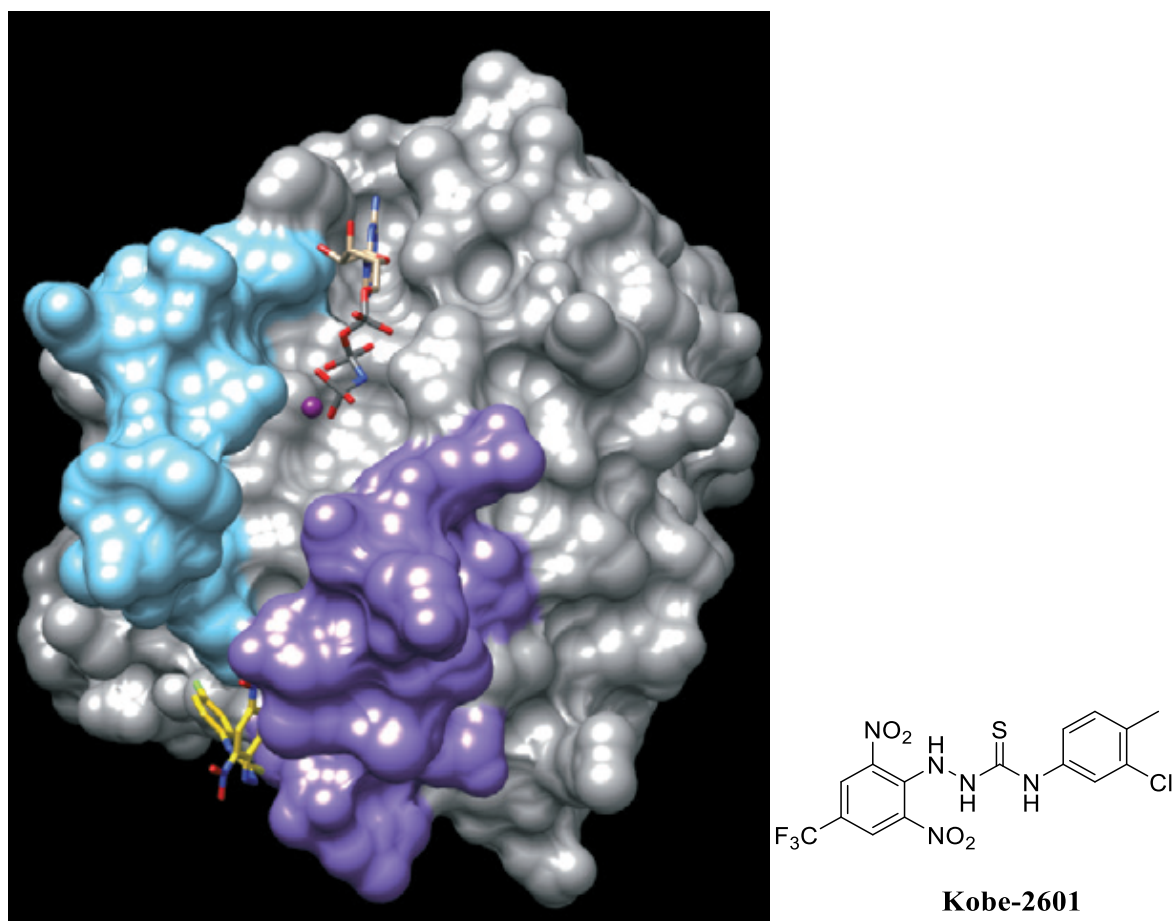


Figure 33; NMR-derived structure of K-Ras<sup>T35S</sup>-GNP bound to **Kobe-2601** (PDB ID: 2LWI)

These compounds also lack potency to be a viable treatment and again are proposed as a scaffold for optimization.

### Mutant-specific inhibitors

In a recent approach, first published by Shokat *et al.*, compounds have been designed that covalently bind the specific mutant K-Ras<sup>G12C</sup>. These inhibitors were discovered by screening a library of small molecules with GDP-bound K-Ras<sup>G12C</sup> using a tethering approach. The most active compound is **Shokat compound 16**. The design and synthesis of analogues was performed with the help of X-ray structures of co-complexes. Crystallographic studies reveal the formation of a new pocket beneath the effector binding switch-II region. Binding of these inhibitors to K-Ras<sup>G12C</sup> disrupts switch-I and switch-II, changing the nucleotide preference to favour GDP over GTP avoiding Raf subsequent binding (Figure 34).

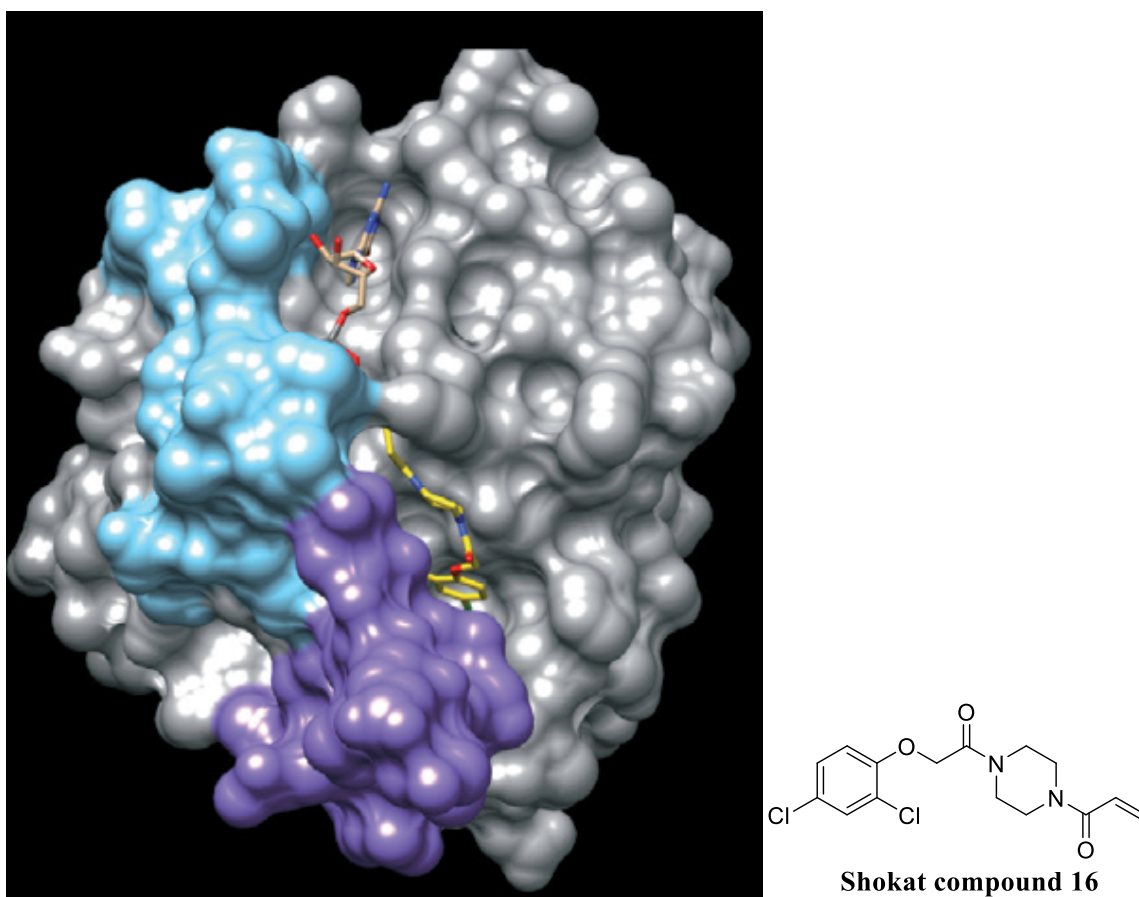


Figure 34; X-Ray structure of K-Ras<sup>G12C</sup>-GDP bound to **Shokat compound 16** (PDB ID: 4M22)<sup>104</sup>

These studies show that a specific mutation can be specifically targeted reducing the toxicity of the therapeutic treatment.

### Ras-GDP induced stabilization

Recently, Gray *et al.* prepared a GDP analogue with an electrophile attached (**SML-8-73-1**-Figure 35) that binds to a cysteine of mutant K-Ras<sup>G12C</sup> in the presence of 1 mM concentrations of GDP and GTP.<sup>105</sup> It was presumed that “GDP-bound Ras-like” conformation would induce an inactive K-Ras conformation. As a proof-of-concept the idea works, but the molecules have many pharmacokinetic issues due to the charged nature of the drug.

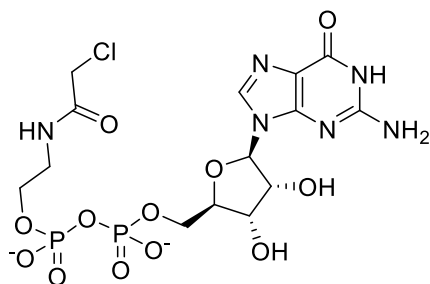


Figure 35; SML-8-73-1

### 2.6.2 New and alternative approaches to target Ras<sup>106</sup>

All the experiences that have been previously described, together with the lack of a well-defined hydrophobic pocket(s), led to the idea that Ras may be not tractable.<sup>107</sup> This contributed to the pursuit of viable therapeutic treatments targeting proteins that promote Ras membrane interaction, effector signaling, synthetic lethality or Ras-dependent metabolism.<sup>108,109</sup>

#### Farnesyl transferase inhibitors (FTIs)

The protein association with the inner part of the membrane is crucial for the Ras oncogenic activity, therefore post-translational modifications that modulate this association could be useful for the development of Ras drugs that disrupt its membrane association.<sup>110</sup> Ras proteins are initially cytosolic; the first post-translational modification<sup>111</sup> is *farnesyltransferase*-catalyzed covalent addition of a farnesyl moiety to a cysteine residue. The second, in the surface of the endoplasmatic reticulum, is a proteolytic removal of three amino-acids by Ras converting enzyme 1 (RCE1) and the last modification is done by *isoprenylcysteine carboxyl methyltransferase* (ICMT), which facilitates a methyl group transfer to the C-terminal amino acid to prevent plasma membrane repulsion. Extra modifications together with the addition of palmitate fatty acid, promote Ras association to the plasma membrane.

An intense search was done during the 1990's to try to disrupt the Ras association to the membrane by inhibiting the process described above in different stages.

Two inhibitors (lonafarnib and tipifarnib) advanced to phase III clinical trials but showed no efficacy, because the protein is able to find alternative ways to be post-translationally modified and associated to the membrane. In recent years the interest for inhibitors of the membrane association has been revisited and one of the most interesting targets identified is

the *prenyl binding protein phosphodiesterase*  $\delta$  (PDE $\delta$ ). A solubilizing factor that facilitates the transport of Ras proteins within the cell through the cytosol.<sup>112</sup>

Initially it was thought that PDE $\delta$  would overcome the weaknesses of FTIs, but the amount of side effects produced from PDE $\delta$  inhibition is too high.

### **Targeting downstream effector signaling**

Blocking effector signaling is one of the most pursued anti-RAS strategies, however, with at least 11 catalytically effector families, the key question is: which effectors to target, and if concurrent inhibition of several effectors is necessary in the probable case that the cell would be able to bypass a single disruption in the signaling cascade.

The first Raf inhibitors (vemurafenib and dabrafenib) are now approved for B-Raf-mutant melanomas. When these inhibitors were evaluated in Ras-mutant cancer cells, a paradoxical activation of ERK signaling was observed. This activity was determined to be driven by mutant Ras-dependent formation of B-Raf - C-Raf heterodimers.<sup>113</sup> Because the first-generation of Raf inhibitors are only B-Raf selective, they paradoxically cause activation of C-Raf and thereby activation of MEK-ERK signaling. This observation led to the development of second-generation pan-Raf inhibitors (LY3009120 and PLX8394), which do not activate the MAPK pathway in the presence of a Ras mutation.<sup>114, 115</sup>

Potent and highly selective inhibitors of MEK1 and MEK2 have also been developed. Unfortunately, these inhibitors have shown limited efficacy in Ras-mutant cancers, in part because of loss of ERK-mediated feedback inhibitory mechanisms.<sup>116</sup>

Although increased ERK signaling drives cancer growth, aberrantly high ERK activation can be growth-inhibitory. Therefore, feedback inhibition mechanisms are initiated by ERK phosphorylation of components upstream of MEK (like C-Raf or B-Raf), dampening flux through the cascade. Consequently, upon MEK inhibition of ERK, increased flux through Raf and MEK can overwhelm the inhibitor block, leading to ERK reactivation. Cancer-associated MEK mutations have also been described that render these cancers insensitive to MEK inhibitors. These unanticipated consequences of MEK inhibition have prompted the development of inhibitors of ERK, with several candidates now under clinical evaluation.

The PI3K-AKT-mTOR pathway serves as a highly complementary role for the Raf-MEK-

ERK cascade. Resistance to ERK- MAPK pathway inhibitors can be mediated by PI3K-AKT-mTOR activation. Inhibition of components of both pathways has shown enhanced antitumor activity in mouse models. However, these combinations have not demonstrated the same activities in clinical trials, in part because of toxicities not seen in mouse studies.<sup>108</sup>

Recently, Rigosertib has been reported to have an unanticipated activity, acting as a Ras mimetic to occupy the Ras-binding site in Raf and Ras association domains of Ras effectors.<sup>117</sup> However, another recent study reported that rigosertib affinity for Ras-binding and Ras association domains may not be sufficient to block Ras effector signaling. Studies showed that the observed reduction in ERK was due to an indirect mechanism of inhibition of Sos and Raf.<sup>118</sup>

### **Synthetic lethal interaction partners of mutant Ras**

Another indirect approach for targeting mutant Ras is based on the concept of synthetic lethality.<sup>119</sup> Synthetic lethal interaction partners for mutant Ras are genes whose functions are essential in Ras-mutant but not wild-type Ras cells.

In 2009, several groups reported the identification of synthetic lethal interactors of mutant Ras, including several protein kinases (*e.g.*, STK33 and TBK1). However, these interactors failed to demonstrate a functional linkage with mutant Ras. To date, the search for synthetic lethal interactions of mutant Ras has been unsuccessful.<sup>120</sup>

### **Targeting Ras-regulated metabolic processes**

One of the 10 hallmarks of cancer is the need for cancer cells to reprogram energy and nutrient metabolism to support their elevated proliferative state.<sup>121</sup> These increased needs are met by reprogramming various metabolic processes to recycle intracellular fuel sources or to scavenge extracellular components.

Oncogenic Ras is known to drive macropinocytosis, an endocytic process in which cells use their plasma membrane to internalize nutrients.<sup>122</sup> There are no pharmacologic inhibitors that selectively block macropinocytosis, but EIPA [5-(N-ethyl-N-isopropyl)amiloride], an inhibitor of  $\text{Na}^+/\text{H}^+$  exchange, was reported to inhibit this process and thereby impair PDAC tumor formation in vivo. Although the elevated macro-pinocytotic activity in PDAC is Ras-dependent, the effector signaling mechanisms driving this activity remain poorly understood.

Another route used by oncogenic Ras-driven cancer cells to generate what is required for cell proliferation is autophagy. During this process, the cell degrades unneeded intracellular components, providing a supply of amino acids, lipids, and nucleic acids. Chloroquine inhibition of lysosomal acidification prevented tumorigenic growth in mice.<sup>123</sup> Chloroquine (Figure 36) is a nonspecific inhibitor of autophagy and can also inhibit macropinocytosis. The related approved drug, hydroxychloroquine, is now under clinical evaluation for PDAC (clinicaltrials.gov identifier, NCT01506973-Phase II studies).

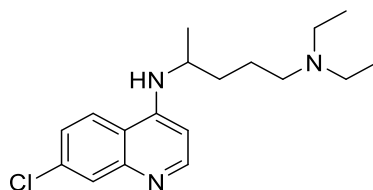


Figure 36; **Chloroquine**

Mutant Ras has been linked to increased glucose metabolism and the use of glucose metabolites into nucleotide and lipid biosynthetic pathways,<sup>124</sup> also Ras-mutant PDAC is characterized by increased dependency on glutamine for NADPH (nicotinamide adenine dinucleotide phosphate, reduced form) production to maintain redox balance<sup>125</sup> but none of these strategies have been studied yet.

## 2.7 Future Perspectives

Despite the challenging history of Ras drug discovery, there is renewed optimism that the goal will be achieved in a reasonably not too far future. It is thought that it will not be a simple therapy that will be effective in all Ras-mutant cancers. The function of mutant Ras is greatly affected by cell type and genetic environment.

### Chapter 3: Project Rationale

In normal cells the GEF activity of Sos is tightly regulated through interactions with Ras, but also by the allosteric modulation of GTP-bound active Ras at an allosteric binding site close to the catalytic Ras binding site present in Sos. Activating mutations in Ras (K-Ras is the most frequently mutated isoform in cancer) short-circuit Sos-mediated activation, as their intrinsic and GTPase-activating protein (GAP)-catalyzed GTP hydrolysis is impaired, leaving them in a perpetually GTP-loaded state.

It was previously assumed that mutant Ras-driven tumorigenesis would be insulated from the effects of inhibition of any up-stream activating factor, including Sos. However, recent evidence suggests that this may be too simplistic an interpretation, and that mutant Ras in fact co-opts wild-type Ras isoforms through allosteric activation of Sos in order to achieve and maintain cellular transformation.<sup>126, 127</sup>

It remains in serious doubt if any of these strategies could yield a potent compound for a potential therapy and we can conclude that the functional output of Ras proteins depends on their GTP-binding status that is dictated by the counteracting actions of guanine nucleotide exchange factors (GEFs) and GTPase activating proteins (GAPs).

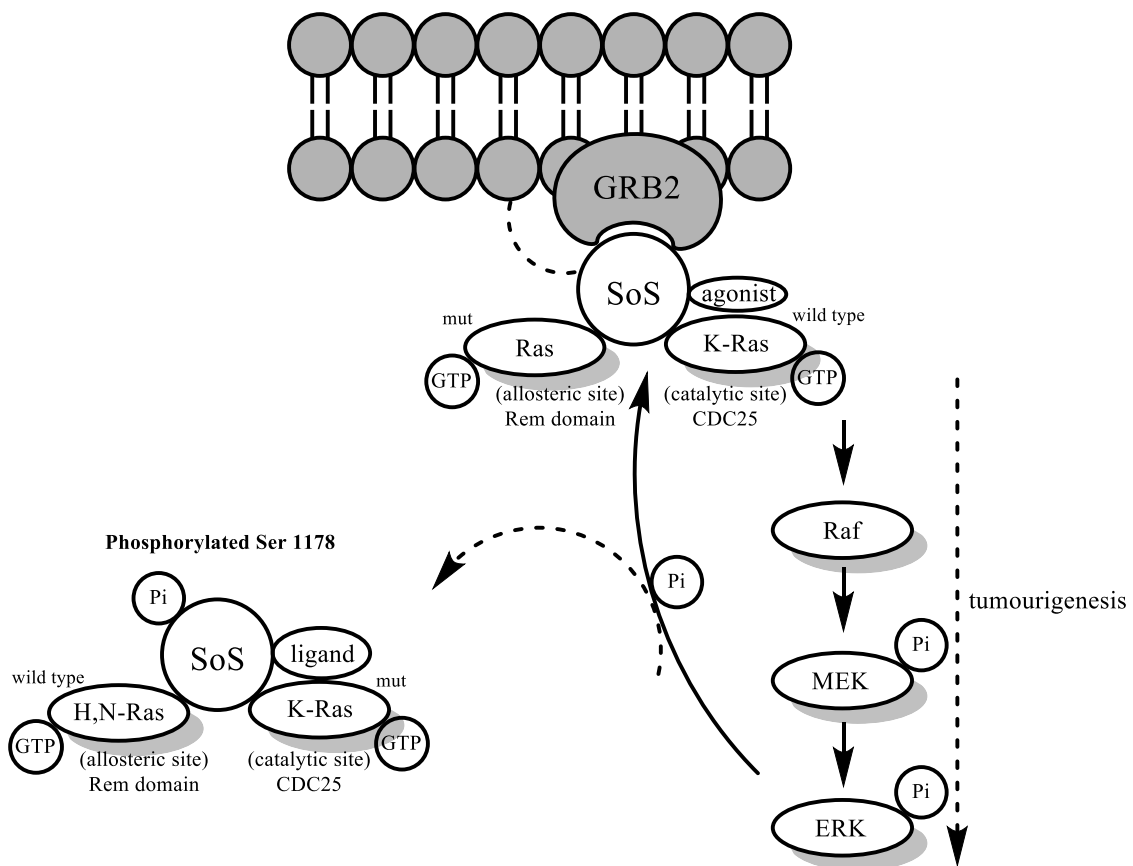
In the absence of oncogenic mutations that compromise the response of Ras to the action of GAPs, the amount of GTP-bound Ras molecules is thought to be controlled by the action of GEFs. The Ras-specific GEF Son of Sevenless, has a key role in regulating Ras activity by catalysing the exchange of GTP for GDP<sup>128</sup>. The presence of an allosteric site on Sos that interacts preferentially with Ras-GTP has been previously reported. This interaction augments the catalytic activity of Sos by affecting conformational changes that release it from its autoinhibited state.<sup>129</sup>

As oncogenic Ras proteins are constitutively bound to GTP, their interaction with the allosteric site on Sos would be favoured. Accordingly, it is plausible that the occupation of the allosteric site on Sos by oncogenic Ras would lead to the enhancement of Sos-mediated activation of wild-type (WT) Ras. Accordingly, with this observation 2 different approaches have emerged:

- 1) **Fesik approach** (Scheme 2): It has been observed by X-Ray crystallography studies that aminopiperidine indole ligands, found after an HTS campaign, bind to the ternary



complex (Ras:Sos:Ras) in a sub-pocket of the CDC25 catalytic site. These ligands act as agonists, therefore inducing high concentrations of Ras-GTP in the cell. This can paradoxically disrupt the signaling cascade acting in a biphasic mode: low concentrations of ligand would induce the production of phosphorylated downstream effectors (MEK, ERK), but a high concentration of ligand can enhance the concentration of phosphorylated ERK up to a point where Sos can be also phosphorylated at a serine residue (Ser-1178), disrupting the interaction with GRB2 and therefore creating a cytosolic phosphorylated quaternary complex which cannot drive tumorigenesis anymore (Scheme 2).<sup>130</sup>



Scheme 2

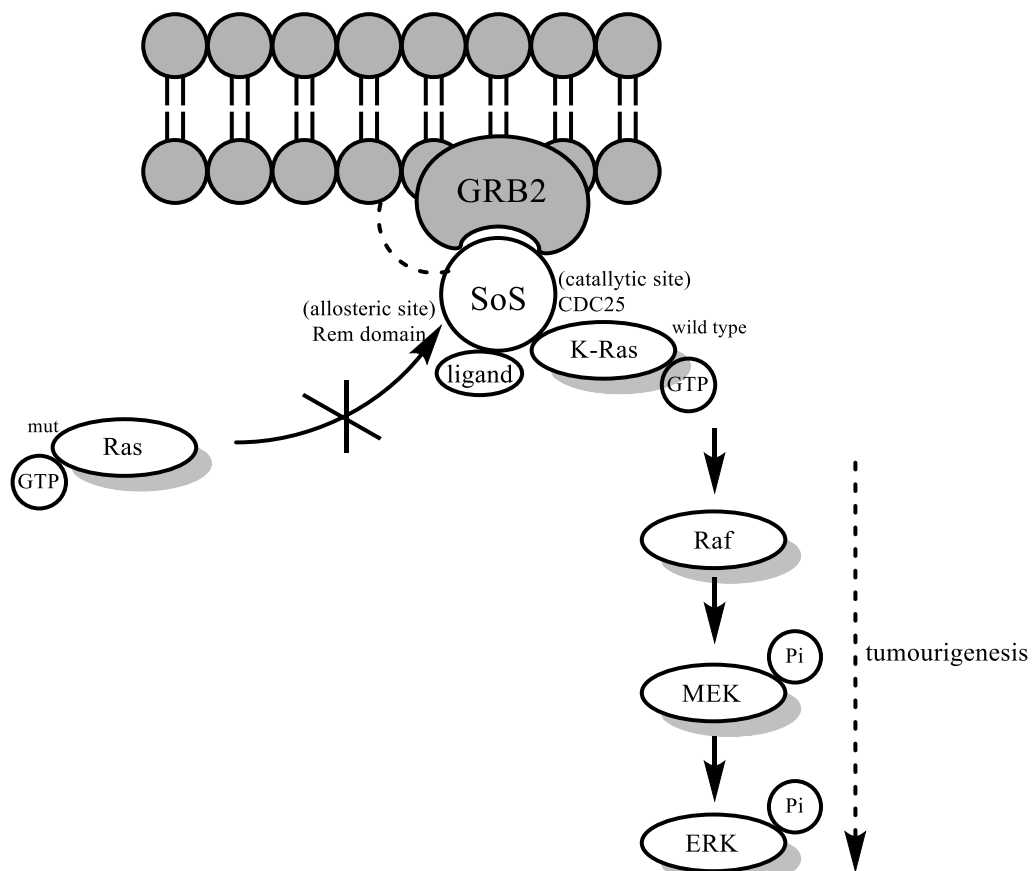
The questions that would remain unanswered with this approach are:

- Why can we not see this phosphorylation happening when tumorigenesis is taking place as phosphorylation is also upregulated?
- Why would the system be stable?

Regarding the latter question it could be theorized that once the concentration of

phosphorylated ERK decreases after disruption of tumorigenesis, the quaternary complex will not form and the cycle would re-start.

- 2) Bar-Sagi approach: On the observation that wild type Ras recruits mutant type Ras (permanently bound to GTP as GAPs activity is impaired) to form a catalytically activated ternary complex (Ras:Sos:Ras), it has been also proposed to disrupt this association by ligands that can expand the cleft between the allosteric site (Rem domain) and the catalytic site (CDC25), then the ability of Sos to recruit mutant Ras will be impaired, and therefore tumorigenesis disrupted (Scheme 3).<sup>88</sup>



Scheme 3

Based on the observations by Bar-Sagi *et al.* that mutant K-Ras is dependent on the Sos-mediated activation of mutant-type Ras, we hypothesize that inhibition of Sos's ability to catalyze nucleotide exchange on wild-type Ras has the potential to target oncogenic Ras signaling.

In preliminary studies, we have identified small molecule compounds that bind to an unprecedented site on Sos itself and have the potential to inhibit Ras activation via a novel

allosteric mechanism (Figure 37).

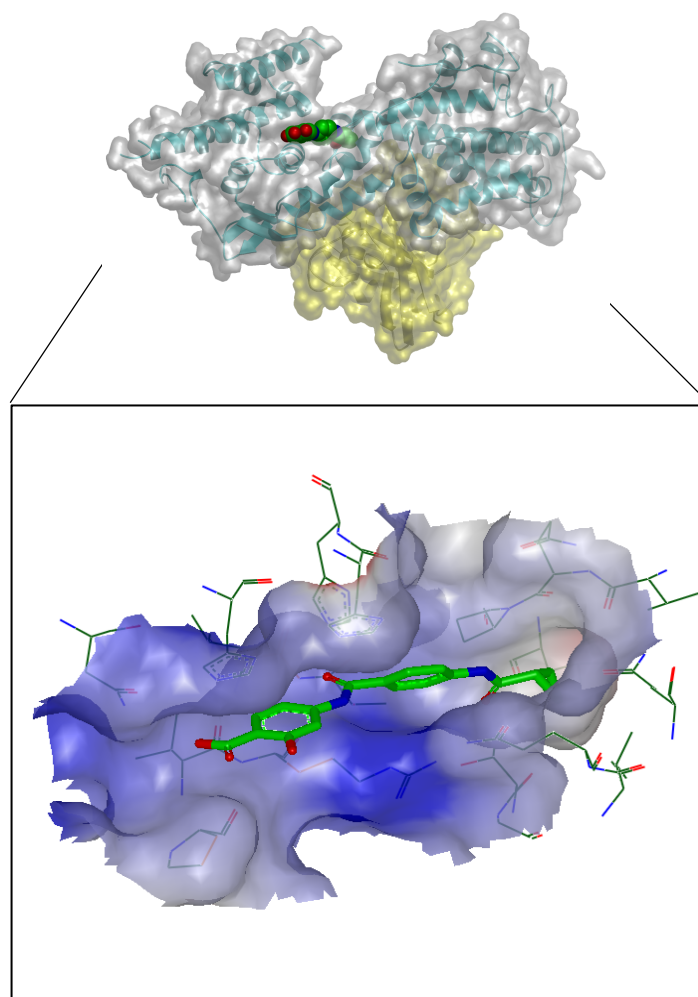


Figure 37, X-ray structure of compound with Sos<sup>cat</sup>:Ras (catalytic site) complex can confirm binding

As we can see in the previous figure, the binding site spans the cleft between CDC25 (where catalytic Ras binds) and Rem domains of Sos (where allosteric Ras binds). We can see a mixed polar/apolar site, giving us opportunities for enhancing the affinity.

We seek to establish this as a viable mechanism to treat Ras mutant pancreatic tumors and pave the way for future drug discovery strategies.

### Preliminary Data

To identify compounds that bind to Sos, we conducted a small-scale biophysical binding screen of a 960-compound library, using a thermal melting ( $T_m$ -shift, thermofluor) assay to detect compounds that would bind to the catalytic domain of Sos (CDC25). From this screen, three compounds were identified as inducing  $T_m$  shift of 1°C or greater (Figure 38).

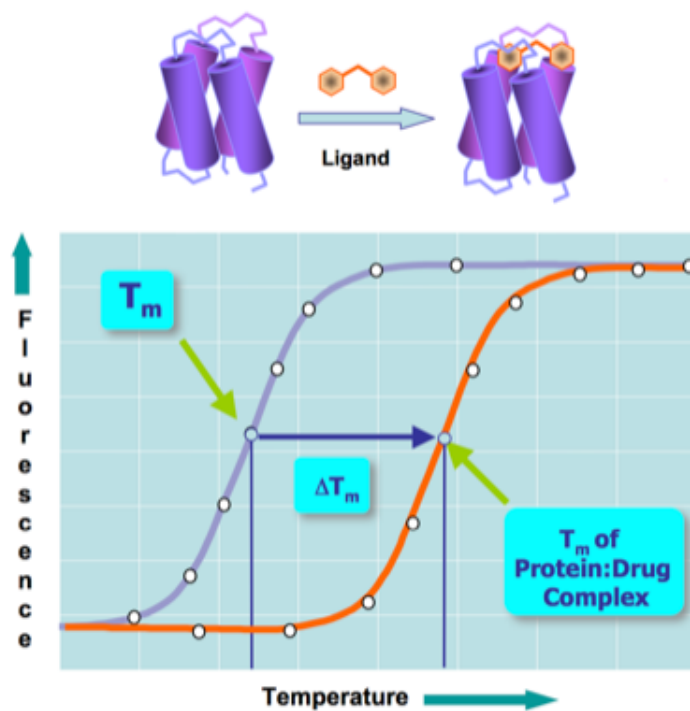


Figure 38; 960-compound library screened against Sos<sup>cat</sup> using 'T<sub>m</sub>-shift' assay

Compound (**11**), that induced a T<sub>m</sub> shift of 1.3°C, was selected for further characterization. The compound was co-crystallized with H-Ras:Sos complex, showing the binding site of (**11**) (Figure 39).

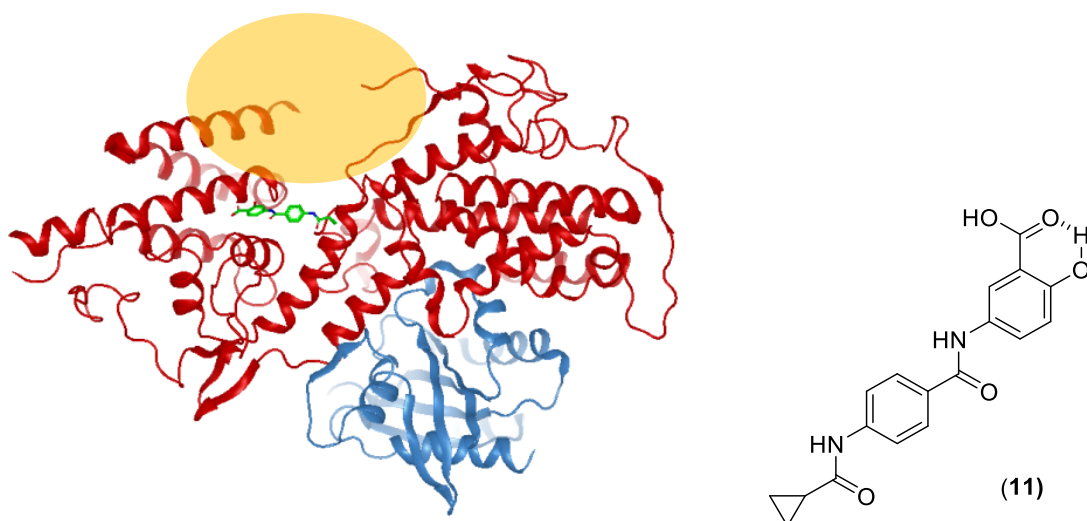


Figure 39; Complex of (**11**) (green) with binary Ras:Sos. Sos<sup>cat</sup> (red), 'catalytic' Ras (blue), Ras allosteric site (orange)

2D protein-observed NMR using  $^{15}\text{N}$ - $^2\text{H}$ -labelled  $\text{Sos}^{\text{cat}}$  confirmed specific binding of **(11)** to Sos and yielded a  $K_d$  of  $\sim 400\ \mu\text{M}$ . Soaking of **(11)** into crystals of  $\text{Sos}^{\text{cat}}$  with H-Ras bound at the catalytic site (Ras:Sos) revealed that the compound binds in a hitherto unreported and unique binding site spanning the “cleft” between CDC25 and Rem domains of Sos. The outer edges of this cleft form the allosteric Ras binding site, via which activated Ras potentiates the catalytic activity of Sos in a positive feedback loop.  $^2\text{D}$ NMR titrations of  $^{15}\text{N}$ ,  $^2\text{H}$ - $\text{Sos}^{\text{cat}}$  previously saturated with H-Ras:GTP $_{\gamma}\text{S}$  (GTP $_{\gamma}\text{S}$ : a non-hydrolyzable analogue of GTP) with **(11)** revealed an apparent  $K_d$  of  $\sim 1.2\ \text{mM}$ , compared with  $400\ \mu\text{M}$  in the absence of Ras. The decrease of measured binding affinity of **(11)** to Sos complexed with Ras at both its allosteric and catalytic sites, supports the hypothesis that **(11)** competes for the binding of Sos allosteric site with Ras. This could be rationalized through comparison of the structure of binary Ras:Sos bound to **(11)** with structures of the ternary complex of Sos with Ras bound at both, the catalytic and allosteric sites (ternary Ras:Sos:Ras). **(11)** would clash sterically with “allosteric” Ras via its salicylate group (Figure 40).

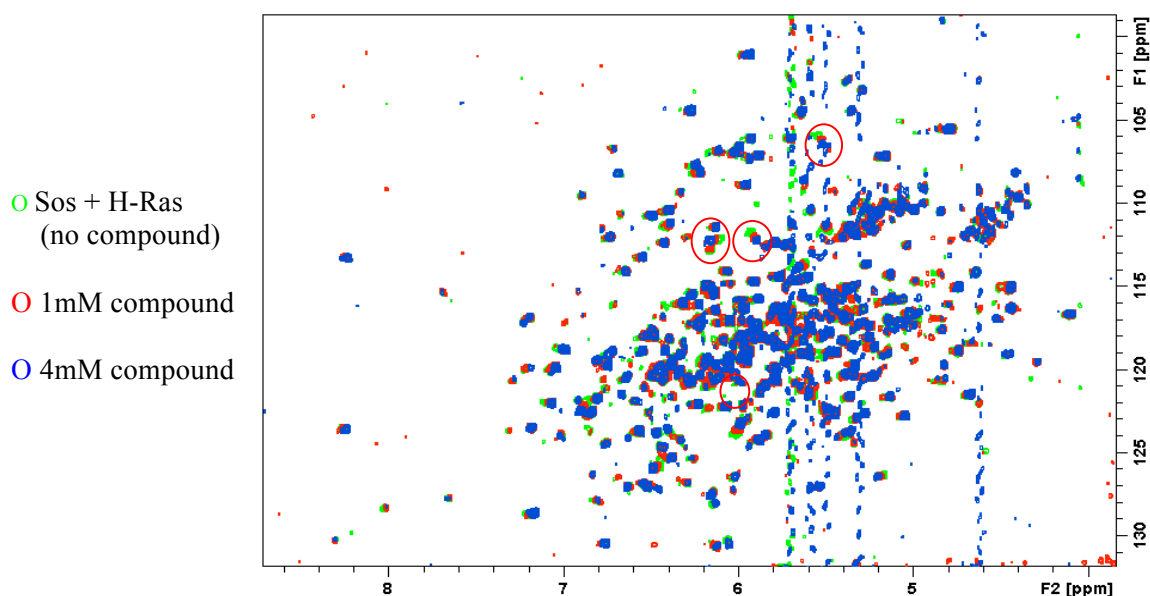


Figure 40;  $^1\text{H}$ - $^{15}\text{N}$  NMR of Sos shows  $K_d$  (affinity) of  $\sim 400\ \mu\text{M}$ , Repeating experiment in presence of GTP-Ras weakens  $K_d$  to  $1.2\ \text{mM}$ , suggesting binding at the allosteric site

The close analogue **(12)** (Figure 41) has also been solved in the binary Ras:Sos complex crystal system and, as expected, adopts the same binding mode; however, the affinity data

suggests a slightly better interaction with a  $K_d$  of 1 mM in the presence of Ras-GTP.

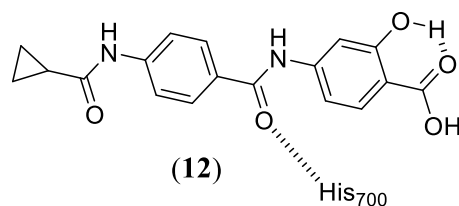


Figure 41

The binding site for (11) and its close analogue (12) is predominantly hydrophobic in nature, and the ligands make only a single polar H-bond contact between the benzoylamino carbonyl and a  $\epsilon$ H-N of His700. However, examination of the site clearly indicates a number of other potential polar interactions that were exploited through relatively minor modifications to the existing scaffold and/or accessed with more significant changes.

### 3.1 Research Plan

The ligand efficiency of the hits we have identified for the novel binding site on Sos is modest (around 0.18 KJ/mol), yet clear potential exists to access additional energetically favorable interactions via appropriate modifications or scaffold changes.

#### 3.1.1 Optimization of compound (11) and (12)

Initially, a small set of analogues of the primary hit (11) was synthesized. We had a number of design concepts that we explored to exploit additional polar and non-polar interaction potential in and adjacent to the binding site of (11). Different biophysical assays that are well established in the Alexander L. Breeze lab ( $T_m$  shift and/or microscale thermophoresis (MST), surface plasmon resonance (SPR) and bi-dimensional  $^1\text{H}$ - $^{15}\text{N}$  TROSY-HSQC and  $^1\text{H}$ - $^{13}\text{C}$ -methyl HMQC-NMR)<sup>131</sup> were used to measure the affinities of these analogues and confirm binding to the allosteric Ras binding site of  $\text{Sos}^{\text{cat}}$ . Initial analogues have been designed based on the existing structural information (Figure 42) and will consist of two components.

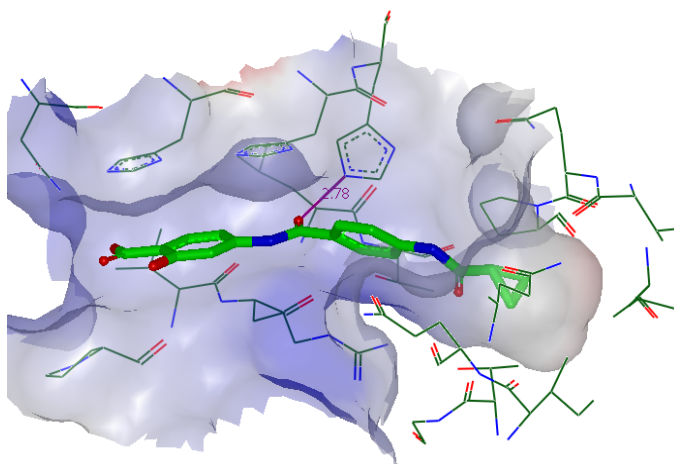


Figure 42; Binding site for **(12)** in  $\text{Sos}^{\text{cat}}$ . The left-hand side is formed by residues from the Rem domain, while a pocket in the CDC25 domain forms the right-hand part of the cavity

Firstly, the existing lead was modified to optimize its contacts and interactions to try to enhance potency by growing the existing structure towards areas of the protein which have the potential for making further productive interactions (Figure 43).

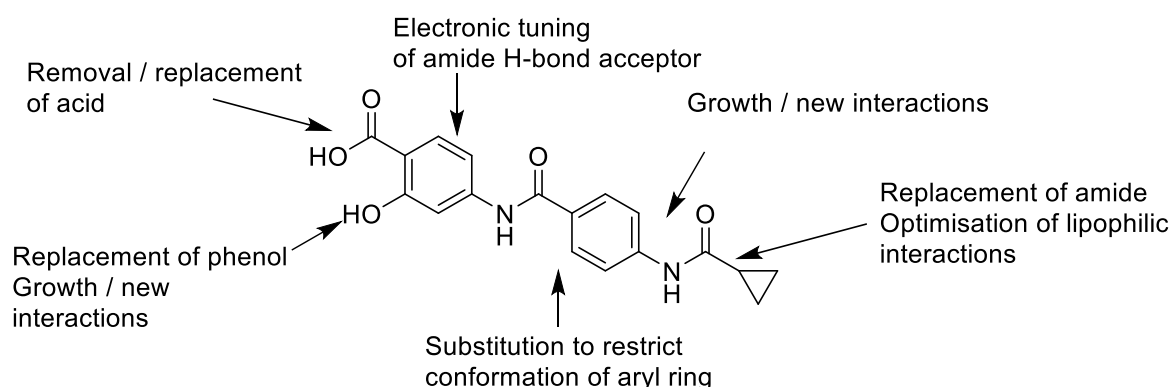


Figure 43; Outline chemical strategy for chemical optimization

It can be seen from the structure that the acidic group does not make interactions with the protein; this is supported by the observation that the acid and phenol motifs can be transposed without significant loss of activity of **(11)** and **(12)**. Addition of N-heterocycles (indazolones, isoindazolones and aminoindazoles) replacing the aminosalicylic group was also proposed.

The existing hydrogen bonding interaction made by the central amide group was tuned electronically by manipulating the electronics of the aryl rings. Structures of this type are likely to suffer from compromised solubility by the formation of strongly hydrogen bonded

arrays in the crystalline state; hence, in order to avoid possible solubility issues, methylation of one or both amides, was explored.

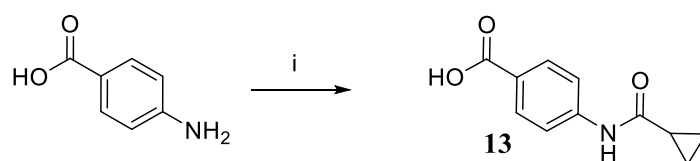
In the crystal structure, the second amide ring can be seen to adopt a non-coplanar configuration with respect to the amide, hence it is likely that potency could be enhanced by the addition of a substituent *ortho*- to the amide.

Finally, the cyclopropyl group occupies a lipophilic pocket within the protein; the fit of the molecule within this pocket was explored by the modification of the cyclopropane ring. The second amide appears not to make productive interactions and so alternative, less polar linking groups were explored. This could increase the potency and provide additional vectors for growing the compounds towards hydrophobic sub-pockets of the protein.



## Chapter 4: Results and Discussion

For the design and synthesis of the analogues, a common intermediate was synthesized and derivatized, this strategy gave us a convergent synthesis and the opportunity to save time and synthetic steps. For this, *para*-amino benzoic acid was treated with  $\text{Na}_2\text{CO}_3$  in dry THF, then cyclopropylcarbonyl chloride was added and the mixture was stirred for 16 hours at room temperature, affording compound (**13**) in 44% yield after purification. Compound (**13**) was used as the common intermediate (Scheme 4).



Scheme 4; i) Cyclopropylcarbonylchloride,  $\text{Na}_2\text{CO}_3$ , THF, r.t., 16h

The scale up of this compound facilitates the access to most of the analogues. The initial exploration was designed for the salicylic acid moiety, which is the terminal phenyl ring. Also, amide methylation was performed for solubility improvement (Figure 44).

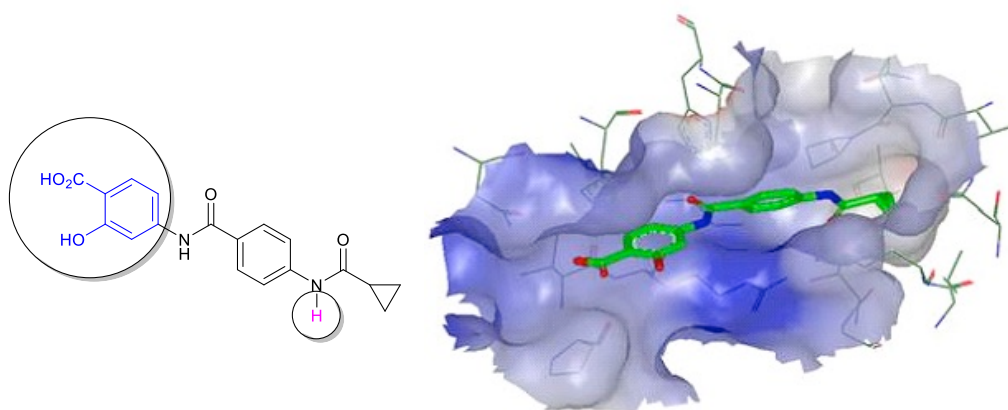
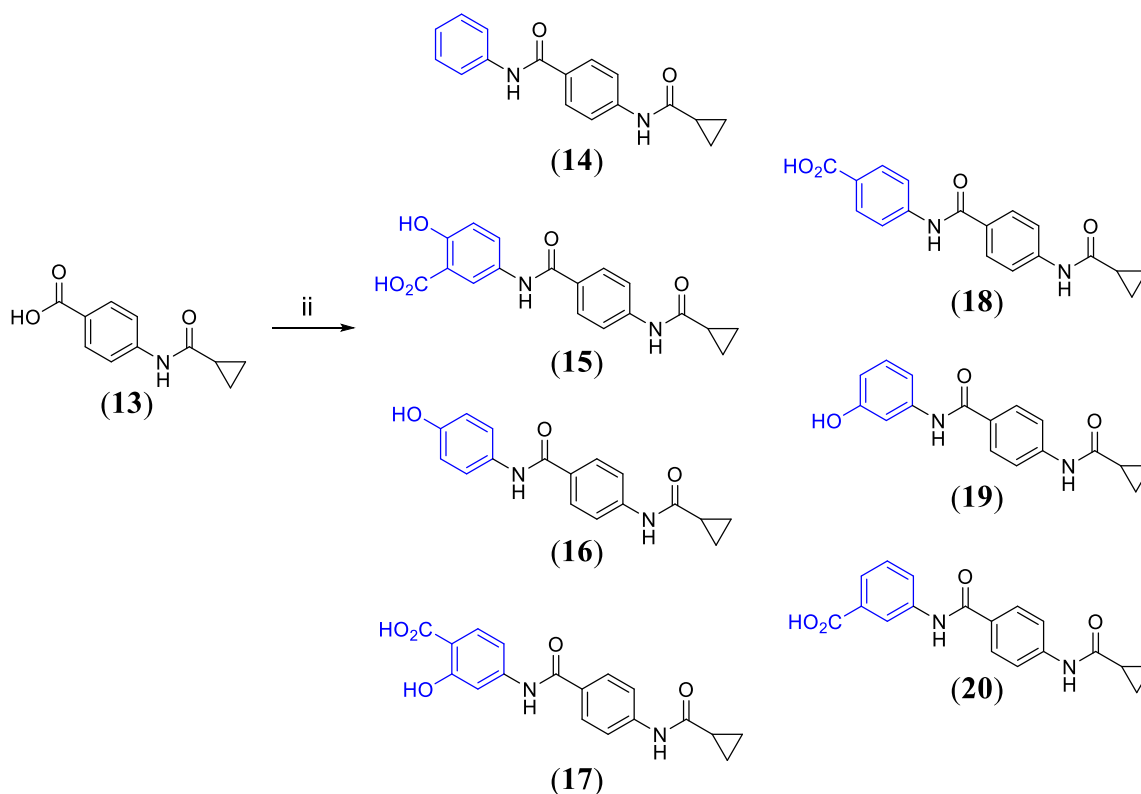


Figure 44; Regions of the molecule of choice for exploration

The exploration performed via the common intermediate (**13**) was carried out to understand the importance of the substituents in the terminal phenyl ring. Elucidation of their most productive position for interactions can help us in building potency. Both substituents or just one of them were removed, leaving the hydroxy or the carboxy groups in positions 4 or 5 of the phenyl ring. For this exploration, several amide couplings were performed using intermediate (**13**) and different anilines. HATU was used as coupling agent after treating

compound (**13**) with DIPEA in dry DMF and with the subsequent addition of the aniline. The reactions stirred for 16 hours at room temperature. The yields obtained were very variable depending on the *para*-substituent of the amino group of the aniline and its electron donating or withdrawing properties (Scheme 5).



Scheme 5; ii) Aniline, HATU, DIPEA, DMF, r.t., 16h

After the synthesis of the analogues shown in the previous scheme, it was noted that the solubility of these compounds was very challenging. Specially for those with a carboxy group in the terminal phenyl ring, which could be purified by washing with DMSO in a sintered funnel. The different analysis performed were done in hot DMSO to avoid precipitation of the analogues.

The solubility issues that we can see in our molecules are ascribed to the polyamide stacking properties of the compounds that we are handling. The amides will interact with each other creating very insoluble solid compounds (Figure 45).

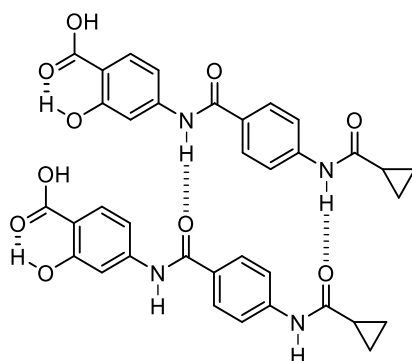
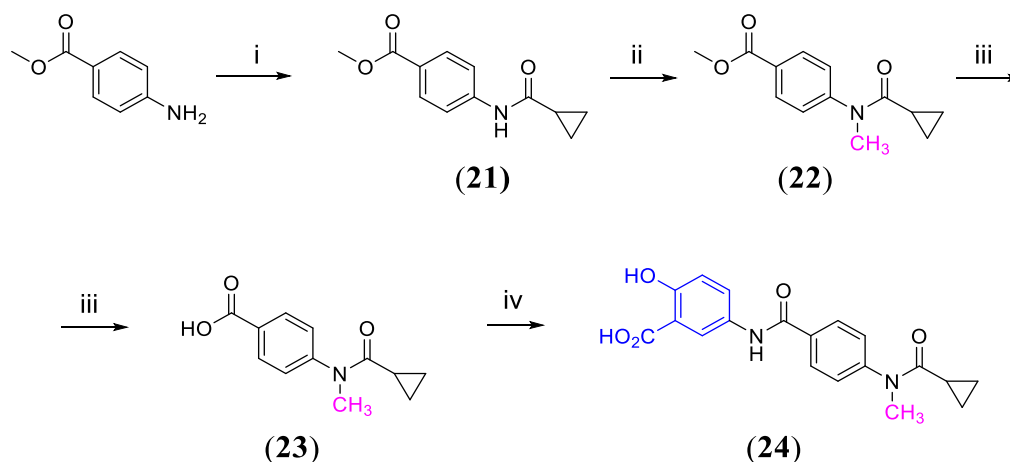


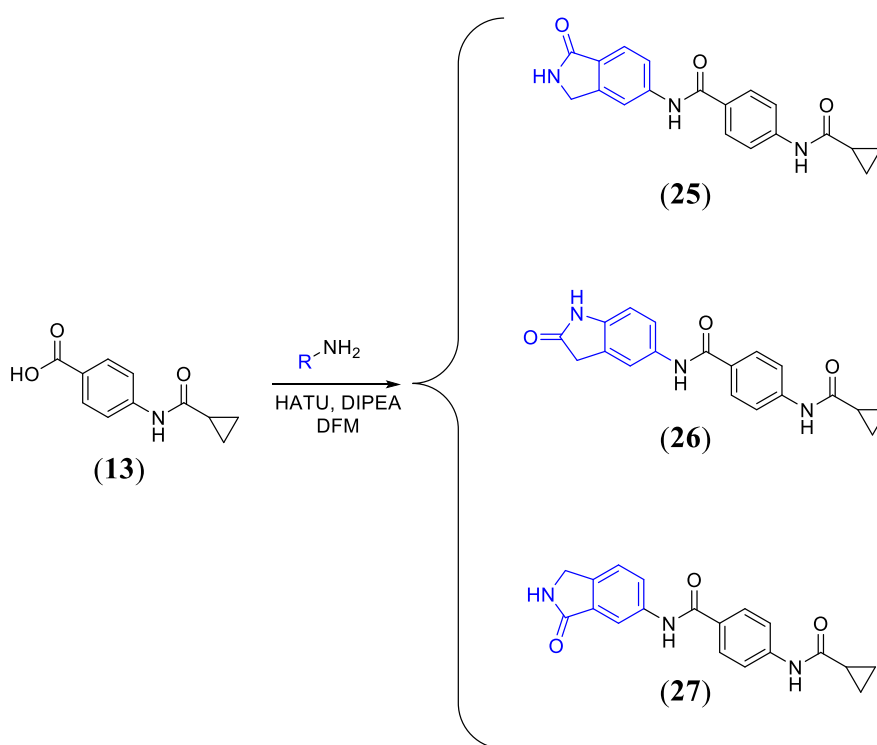
Figure 45; Putative H-bonding repeat interactions

It was theorized that the methylation of one of the amides could disrupt the interaction of the carbonyl group of one molecule with the hydrogen of the amide of the next. To synthesize the new derivative a new synthetic route was designed. To avoid methylation of the acid present in the terminal phenyl ring, the pathway was started with methyl 4-aminobenzoate. The aniline was reacted with cyclopropylcarbonyl chloride after treatment with  $\text{Na}_2\text{CO}_3$  in dry THF. The reaction was stirred for 16 hours at room temperature affording the desired intermediate in 95% yield. Then, compound **(21)** was treated with sodium hydride (60% in mineral oil) in dry THF at  $0^\circ\text{C}$  and subsequently reacted with iodomethane to afford compound **(22)** after 16 hours, allowing the reaction to warm to room temperature and affording compound **(23)** in 95% yield. The methylated intermediate **(23)** was hydrolyzed under basic conditions yielding acid **(23)** in quantitative yield. Finally, an amide coupling of intermediate **(23)** with 5-amino-2-hydroxybenzoic acid gave us the desired product **(24)** (Scheme 6).



Scheme 6; i) Cyclopropylcarbonyl chloride,  $\text{Na}_2\text{CO}_3$ , THF, r.t., 16h; ii) Iodomethane, NaH (60%), THF, r.t., 16h; iii) 2M NaOH,  $\text{H}_2\text{O}$ , r.t., 16 h; iv) Aniline, HATU, DIPEA, DMF, r.t., 16h

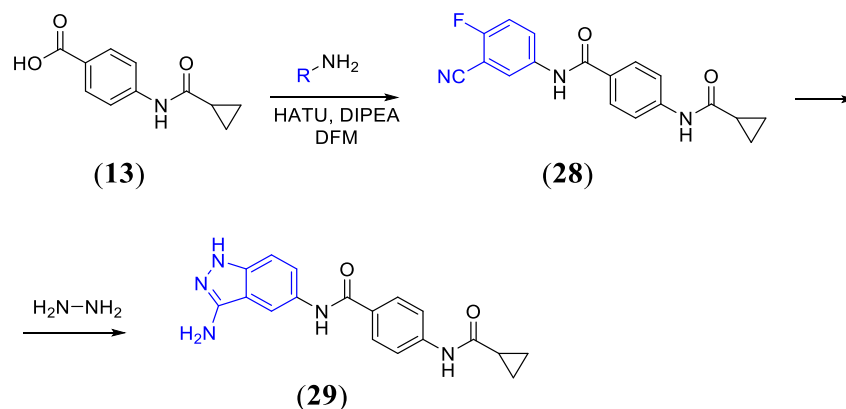
Additionally, the presence of a new heterocycle derivatives on the terminal phenyl ring could give an improvement in potency if the molecule fits in the appropriate way in the pocket. This approach gave us also the opportunity to try different functional groups. Indolinones, isoindolinones and aminoindazoles were synthesized. For the synthesis of indolinones and the isoindolinone, similar chemistry to the one shown in previous schemes was used. Intermediate **(13)** was reacted with the appropriate 4-amino isoindolinone/4-amino indolinone after treating compound **(13)** with HATU in the presence of DIPEA in dry DMF (Scheme 7). The yields of these amide couplings slightly improved (31-52%) after changing the solvent to DMF, which was crucial for better solubility of the aminoisoindolinone and aminoindolinones.



Scheme 7; Amide coupling reaction for the derivatization of a common intermediate

We were also interested in the synthesis of aminoindazoles as these molecules also had a hydrogen bond acceptor mimicking the carboxylate group. The synthesis of these compounds was completely different to the synthesis shown in previous schemes. A fluoronitrile intermediate was used to ring close the aminoindazoles using hydrazine hydrate. 4-amino-2-fluorobenzonitrile was very deactivated due to presence of a nitrile in *para* position to the amino group and the amide coupling was not successful. In the case of 5-amino-2-fluorobenzonitrile the amide coupling was successful but with a low yield (52%). Compound

(**28**) showed multiplicities in the  $^{13}\text{C}$  resonances due to coupling to  $^{19}\text{F}$ . This intermediate was also used for testing. Part of intermediate (**28**) was treated with hydrazine monohydrate at  $60^\circ\text{C}$  for 24 hours yielding the desired aminoindazole (**29**) in 39% yield (Scheme 8).



Scheme 8; Different strategies for the synthesis of aminoindazoles

An exploration with MoE of different compounds allowed us to find a potential new interaction from an aminomethyl moiety in two different ways with two different glutamine residues. On one side, the nitrogen of the amino group is acting as a hydrogen bond acceptor with the hydrogens of the carbamate moiety of a glutamine, and on the other side the same  $\text{NH}_2$  from our ligand is acting as a hydrogen bond donor with a carbonyl of an amide moiety of a different glutamine (Figure 46 – blue circle).

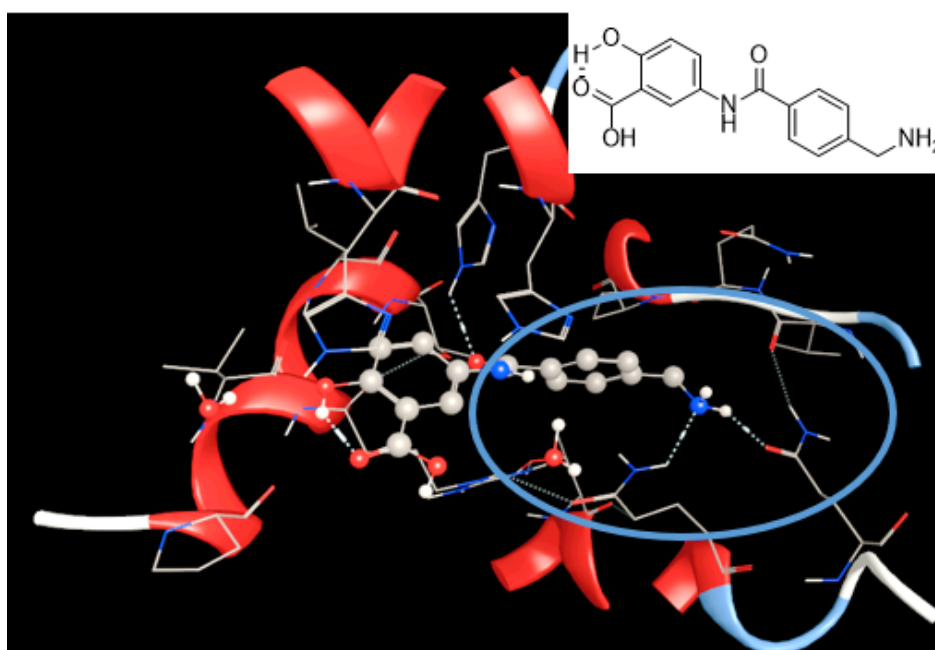
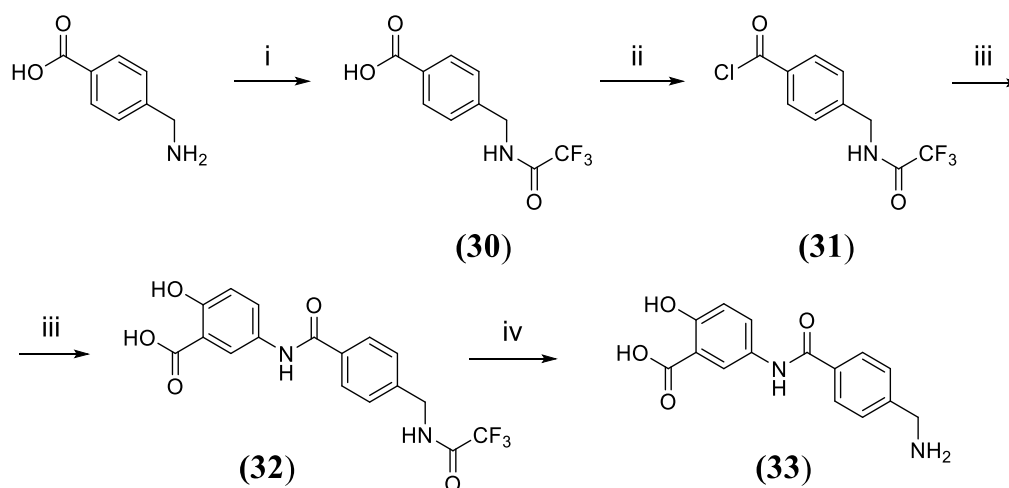


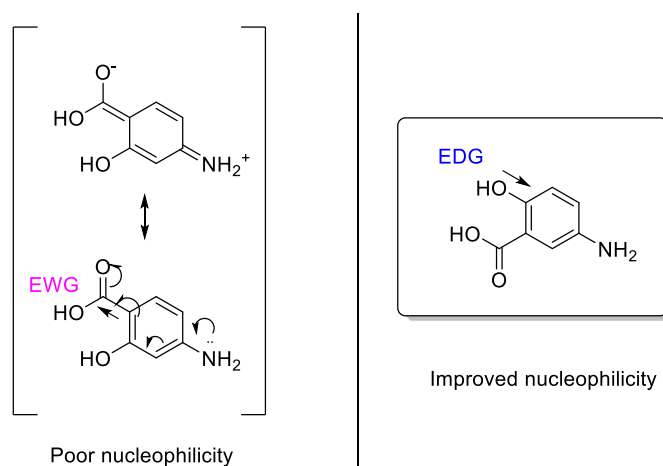
Figure 46; Putative amino group interactions highlighted (PDB ID: 3RSJS)

For the synthesis of the new target, the amino group of the 4-aminomethyl-benzoic acid was protected. The reaction was performed with trifluoroacetic anhydride in DCM at 0°C for 16 hours to give us compound **(30)**. Then the protected intermediate crude product was treated with thionyl chloride to give the acyl chloride **(31)**, which was reacted with 5-amino-2-hydroxybenzoic acid in dry THF for 16 hours at room temperature yielding compound **(32)** in 65% yield over three steps. Subsequent deprotection of compound **(32)** in a mixture of NH<sub>3</sub>/H<sub>2</sub>O (1:9) at 80°C for 16 hours gave us the desired product **(33)** (Scheme 9).



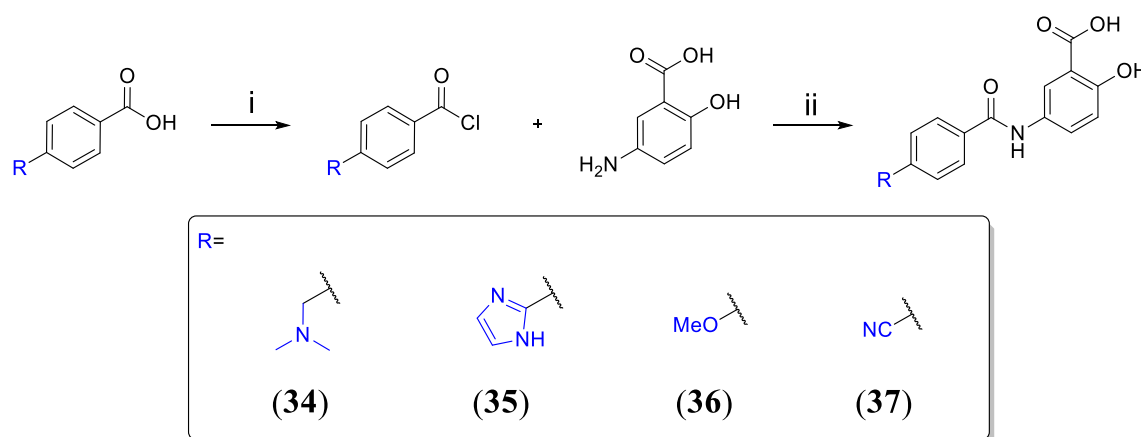
Scheme 9; i) (CF<sub>3</sub>CO)<sub>2</sub>O, pyridine, DCM, 0°C to r.t., 16h; ii) SOCl<sub>2</sub>, 70°C, 16h; iii) 2-hydroxy-4-amino carboxylic acid, THF, r.t., 16h; iv) NH<sub>3</sub>:H<sub>2</sub>O (1:9), Dioxane, 90°C, 16h

In the case of the 2-hydroxy-5-aminobenzoic regioisomer, the lack of aniline nucleophilicity resulted in extremely poor yields, this lack of reactivity led us to focus our efforts on the 2-hydroxy-5-aminobenzoic regioisomer (Scheme 10) for future analogue series.



Scheme 10

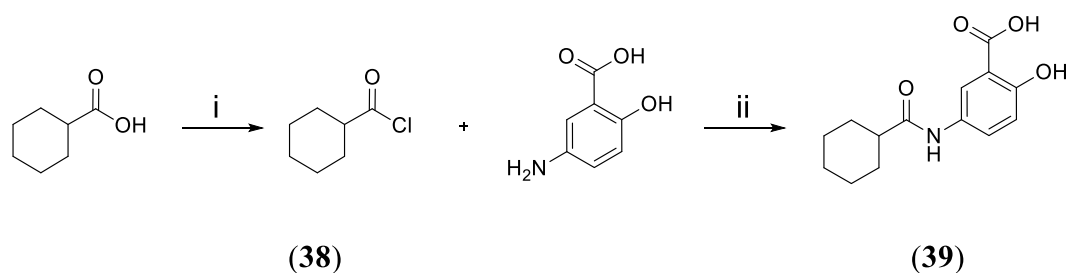
Unfortunately, compound (**33**) lacked enough solubility for biological testing. NMR spectra were obtained by adding TFA to improve the solubility. However, it was decided to synthesize a series of analogues of this compound to try to explore several substituents in the same position of the compound. All the compounds were synthesized by preparing the acyl chloride with thionyl chloride and reacting the crude intermediate after evaporation with 4-aminosalicylic acid in dry THF for 16 hours (Scheme 11). The synthesis of these intermediates was done by Shaun R. Stevens and me.



Scheme 11; i)  $\text{SOCl}_2$ ,  $70^\circ\text{C}$ , 16h; ii) 4-aminosalicylic acid, THF, r.t., 16h

As mentioned before, the solubility of the molecules was very problematic and in some cases the molecules synthesized could not be tested. For this reason, it was decided to make further modifications of the parent compound trying not to lose (or improve) the potency and solubility.

To make the compounds more soluble we attempted to disrupt the hydrogen bonding capacity of the compounds by methylation of the amide bonds as done before with compound (**24**) (Scheme 8), or introducing greater 3 dimensionality,<sup>132</sup> leading to low packaging energies. For this purpose, it was proposed to saturate one of the benzene rings. Accordingly, analogues in which the terminal phenyl ring was replaced with cyclohexyl moiety were explored. The synthesis procedure was similar to that used for previous compounds. After preparation of the acyl chloride with thionyl chloride at  $80^\circ\text{C}$ , the crude intermediate (**38**) was reacted with 4-aminosalicylic acid in dry THF at room temperature for 16 hours, affording compound (**39**) (Scheme 12).



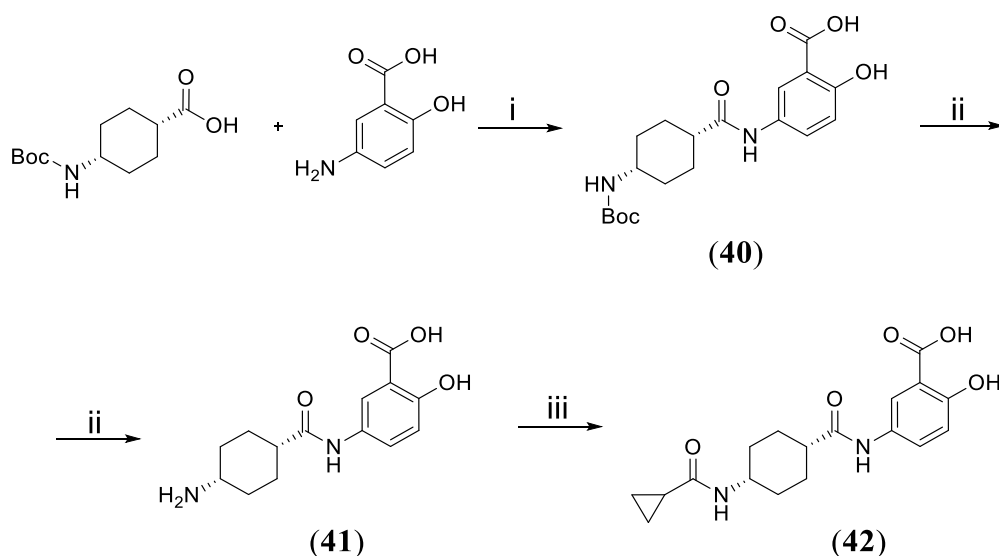
Scheme 12; i) SOCl<sub>2</sub> (solvent), 80°C, 16h; ii) THF, r.t. 16h

Gratifyingly, the solubility of this compound improved compared to previous analogues. Compound **(39)** was soluble in methanol and DMF, which made the analysis of this compound easier.

This information encouraged us to start working with this structure and to synthesize new analogues based on the cyclohexyl moiety. To understand how our new molecules would bind to the designated pocket, different isomers were synthesized exploring the cis- (Scheme 13) and trans- isomers (see Scheme 14). For both isomers duplicated signals due to presence of chair conformers were observed in the NMR analysis (see general experimental).

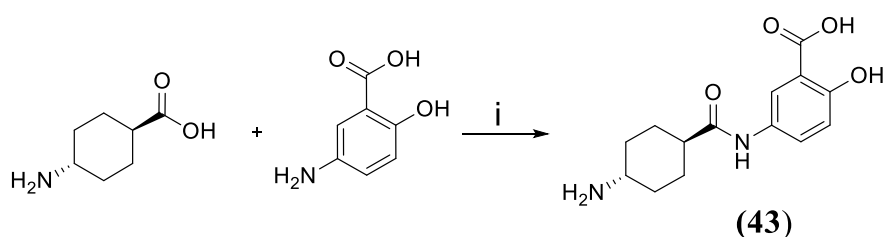
The synthetic procedure started with an amide coupling of the Boc-protected aminocyclohexanoic acid with 4-aminosalicylic acid after treatment with HATU in the presence of DIPEA in dry THF. After 16 hours stirring at room temperature, intermediate **(40)** was obtained in 8% yield. Then, this intermediate was treated with 2M HCl in THF at room temperature for 16 hours affording compound **(41)** in quantitative yield. The deprotected intermediate was then reacted with cyclopropylcarbonyl chloride in dry THF for 16 hours, yielding compound **(42)** in 24 % yield.





Scheme 13; i) HATU, DIPEA, THF, r.t., 16h; ii) aq. 2M HCl, THF, r.t., 16h; iii) cyclopropylcarbonyl chloride, THF, r.t., 16h

In the case of the *trans*- isomer, the Boc-protected version of the starting material was not commercially available, and we decided to try to start the synthesis directly with the deprotected 4-aminocyclohexanoic acid. The amide coupling with 4-aminosalicylic acid was performed with HATU as a coupling agent yielding intermediate (43). This compound has solubility issues due to the high polarity of the molecule and its zwitterionic properties (Scheme 14).

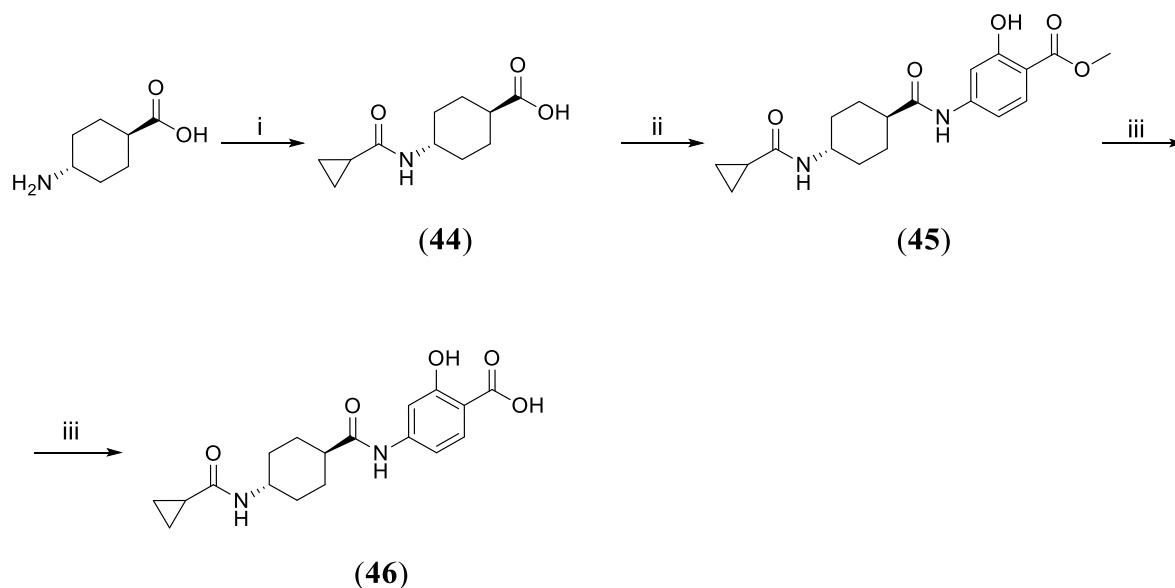


Scheme 14; i) HATU, DIPEA, THF, r.t., 16h

To try to amend this problem a new synthetic route was proposed to decrease polarity and obtain a compound more likely to be soluble in organic solvents. The protection of the carboxylate group of (45) avoids the formation of a zwitterionic compound.

For this, 4-aminocyclohexanoic acid was treated with HBTU in the presence of TEA in dry THF to then perform an amide coupling with cyclopropane carboxylic acid to give compound (44) in 23% yield. This intermediate was treated with EDCI.HCl in the presence of DIPEA in

dry DMF and reacted with methyl 5-amino-2-hydroxybenzoate to afford compound **(45)** in 52% yield. Then, compound **(45)** was hydrolyzed under basic conditions to give the desired product **(46)** in 92% yield (Scheme 15).



Scheme 15; i) Cyclopropanecarboxylic acid, HBTU, TEA, THF, r.t., 6h; ii) EDCI.HCl, DMAP, DMF, methyl 5-amino-2-hydroxybenzoate, r.t., 48h; iii) 2M NaOH, THF, 80°C, 5h

Unfortunately, the solubility of the *cis*- and *trans*- isomers was not as improved as we hoped. Despite being soluble in methanol and DMF in the laboratory, the biological assays were not possible according to the biologists and new approaches were explored.

To try to disrupt the planarity of our molecules and then improve the solubility, another reported approach is the modification of the dihedral angle involved in the amide couplings. For this purpose, introduction of substituents in *ortho*-position to the carboxamide was explored by Harry Shrives (Figure 47).

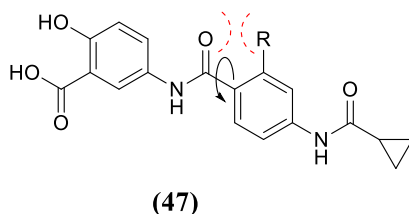


Figure 47; Alkyl steric hindrance inducing orthogonal structure (R= O-Me)

We were happy to see that a methoxy substituent improved significantly the solubility of the compound. At this stage it was theorized that 2 different events could be happening in our molecule to favor solubility:

- The steric hindrance between the methoxy group and the carbonyl group of the amide favors twisting the central aromatic ring out of the plane of the amide (Figure 48-a).
- It could be theorized that our compound is forming a new hydrogen bond. The lone pair of the oxygen and the hydrogen of the amide bond interact, and this hydrogen bond disrupts the polyamide structure and favors solubility (Figure 48-b).

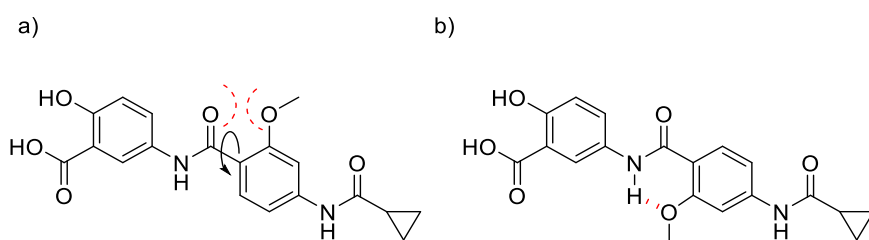


Figure 48; a) Orthogonal-inducing effect by steric hindrance; b) Hydrogen bonding by methoxy group

To confirm which event is enhancing the solubility of our compound we thought that the methylation in ortho- position of the carboxamide (Figure 49) should also twist the central aromatic ring out of the plane. The absence of a hydrogen bond acceptor would discard the “b) approach” as responsible for the improvement of solubility, at least in this particular case.

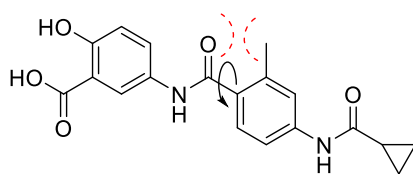
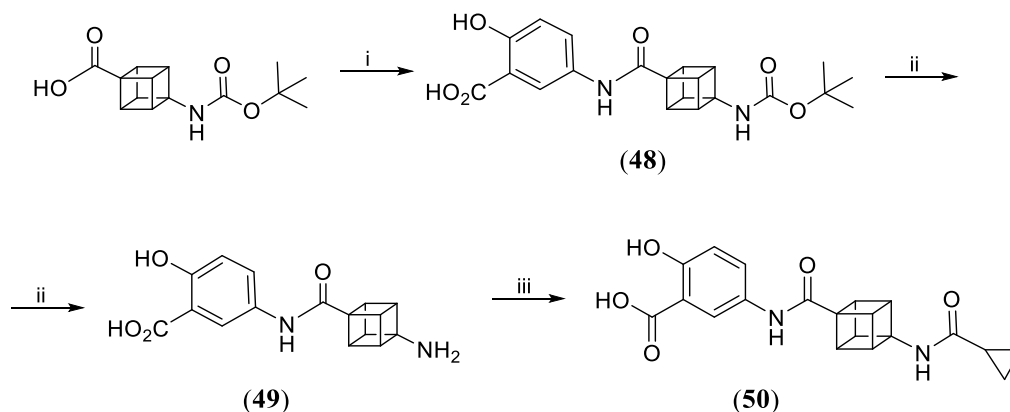


Figure 49; Rotation of the dihedral angle by steric hindrance of the alkyl group

For further improvements of solubility and disruption of the polyamide structure, it was also theorized that the replacement of the central benzyl ring could be positive. To achieve a three-dimensional structure the synthesis of a cubane derivative was proposed. (1s,2R,3s,4r,5S,6r,7R,8S)-4-((tert-butoxycarbonyl)amino)cubane-1-carboxylic acid was reacted with 4-aminosalicylic acid after treatment with HATU in the presence of DIPEA in dry DMF for 16 hours at room temperature, yielding the desired intermediate in 20% yield.

Then, compound (**48**) was treated with TFA in DCM for 4 hours at room temperature, affording the Boc-deprotected compound (**49**). This intermediate was treated with TFA in DCM at room temperature for 4 hours affording compound (**49**). The intermediate was reacted with cyclopropylcarbonyl chloride in dry THF, the mixture stirred at room temperature for 16 hours yielding the desired compound (**50**) (Scheme 16).



Scheme 16; i) 5-Aminosalicylic acid, TEA, HATU, DMF, r.t., 16h; ii) TFA, DCM, r.t., 4h; iii) Cyclopropylcarbonyl Chloride, TEA, THF, r.t. 16h

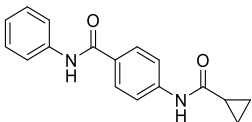
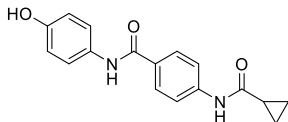
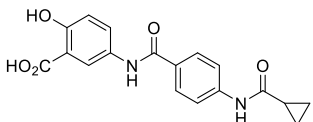
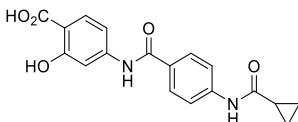
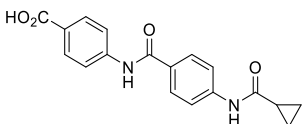
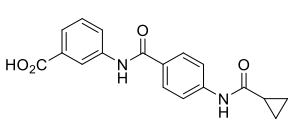
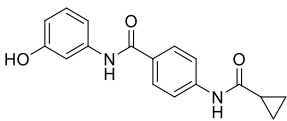
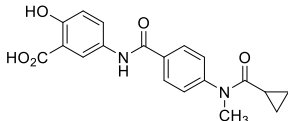
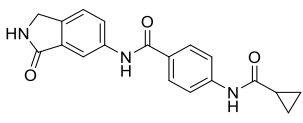
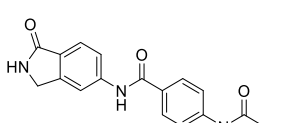
Regarding the solubility observed in the laboratory, the studies performed indicated that the disruption of the polyamide  $\pi$ -stacking can improved the solubility of our compounds by following four different strategies:

1. The methylation of one of the amides also disrupts the polyamide stacking improving solubility (Scheme 8 – Compound (**24**))
2. The introduction of a three-dimensional moiety was achieved by the saturation of one of the benzene rings (Scheme 12, 13, 15 – Compound (**39**), Compound (**42**), Compound (**46**)) or by the introduction of a cubane (Scheme 16 – Compound (**50**))
3. Introduction of an ortho-substituent (Figure 45 – Compound (**47**)) in the terminal phenyl ring close to the ketone, twisting the aromatic ring out of the amide plane. In the case of introducing a methoxy group the same event might be occurring, but hydrogen bonding by the oxygen's lone pair with the amide's hydrogen is also possible (Figure 44)

4. Removal of the right-hand side amide bond and introduction of different substituents also improved solubility (Scheme 13 – Compound (**36**), Compound (**37**))

## Chapter 5: Biological Assays

Unfortunately, the improvements observed in solubility in the laboratory, showing promising solubility in methanol and/or DMF, were not always confirmed by the biologists during the different biophysical assays. In most of the cases, the compounds precipitated during the biological assays under the conditions used: Tris 100 mM, sodium formate 4.5 M, 5% glycerol, pH 8; or Hepes 100 mM, Sodium Formate 4.5 M, 5% glycerol, pH 7.5. Testing our compounds to direct the improvement in potency of the inhibitors was very challenging. Several biophysical studies were trialed by Sebastien Cardon at Leeds University (under the supervision of Alexander Breeze) yielding different outcomes. The solubility of the different compounds synthesized in different biophysical assays is summarized below (Table 1).

ID	Structure	-	ID	Structure	-
14			16		
15			17		
18			20		
19			24		
27			25		

26			28		
29			33		
39			42		
46			34		
35			36		
37			47		
50					

Table 1; Solubility of the compounds in the assays trialed (soluble, partially soluble, not soluble)

- Protein X-Ray crystallography:

To monitor and characterize the protein-ligand interactions by X-ray crystallography the protein complex H-Ras:Sos<sup>cat</sup> was produced, purified and crystalized. The complex was

expressed in mycobacterium tuberculosis (TB) at 37°C. Protein expression was induced by isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) (180 rpm/ overnight).

After the purification of the complex several conditions were trialled, making variations on the concentration of sodium formate and pH conditions to afford crystals (Figure 50).

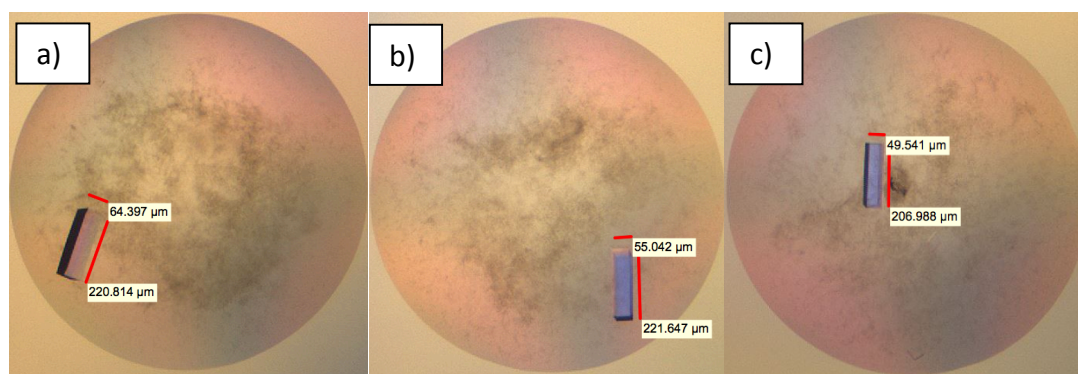


Figure 50; a) 12 mg/ml; b) 10 mg/ml; c) 8 mg/ml

Of all the compounds synthesized, only the hit compounds were soluble and some of them partially soluble ((**33**) and (**47**)) in the soaking conditions (Tris 100 mM, sodium formate 4.5 M, 5% glycerol, pH 8). Unfortunately soaking the compounds in crystal for X-Ray diffraction of the protein-ligand complex cannot guide us in the search of a more potent compound due to insolubility issues.

- Protein Observed-NMR assay:

For protein observed NMR the expression of labelled  $^{2D,15N} \text{Sos}^{\text{cat}}$  is necessary. The experiments were performed in 250 mL of  $\text{D}_2\text{O}$  + 5 g/L of  $^{2D,15N}$  celtone powder. The preculture of the protein was expressed in TB and then treated with labelled media in  $\text{D}_2\text{O}$ . Finally, the protein was purified by  $\text{Ni}^{2+}$  NTA affinity column. TROSY-NMR experiments were performed using a frequency of 950 MHz by increasing gradually the ligand concentration (300  $\mu\text{M}$ , 1 mM, 3 mM, 5 mM). The observation of NMR shifted signals in the aminoacids of the allosteric site gave us information to extrapolate  $K_d$ 's for the different ligands. However, the solubility of compounds was still a liability and only a few ligands could be analysed. The results obtained are summarized below in Figure 51:



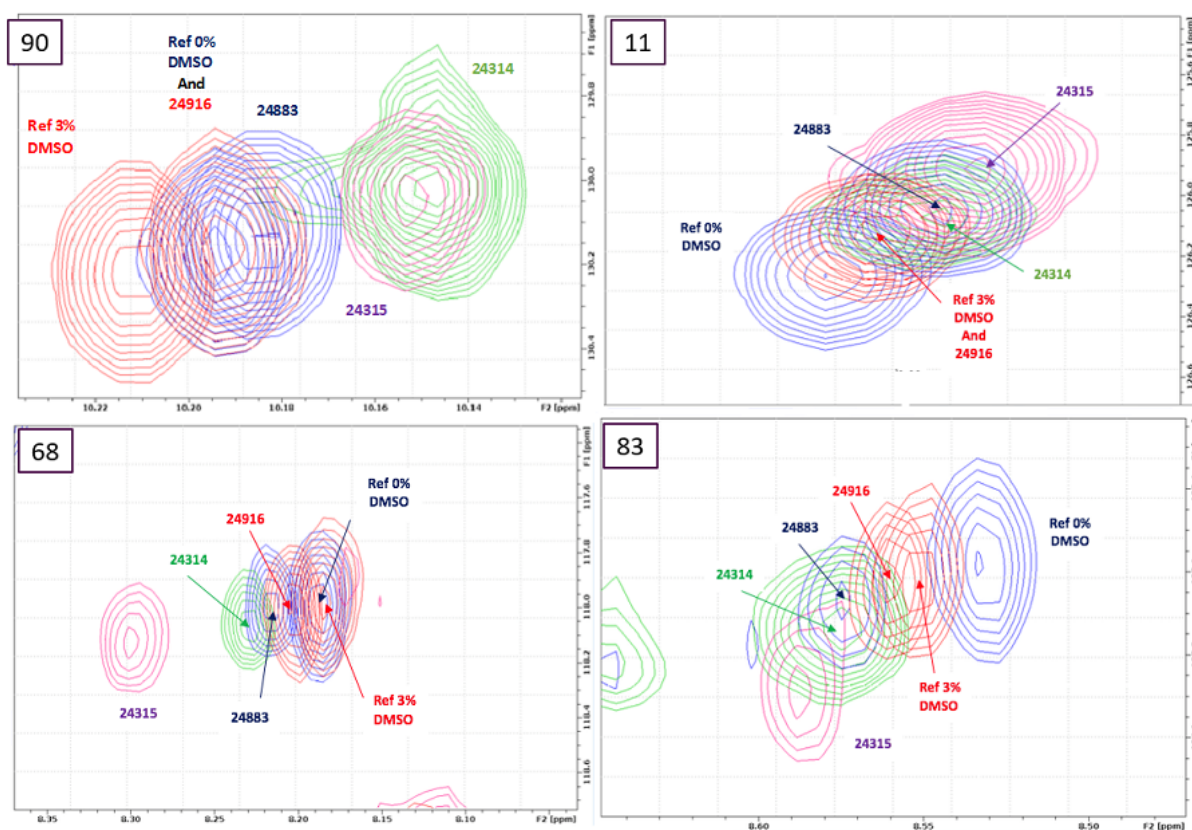


Figure 51; NMR shifts of the  $^{2D,15N}\text{Sos}^{\text{cat}}$  binding site (amino-acids 11, 68, 83 and 90) for compounds Ref. 24314-Compound (15), Ref. 24315-Compound (17), Ref. 24883-Compound (36) and Ref. 24916-Compound (18)

The NMR shifted signals can be observed in several of the allosteric site amino-acids, the same information can be also pictured with all the amino-acids present and extrapolating  $K_d$ 's for each of the most perturbed signal, the information is summarized in Figure 52, 53, 54 and 55:

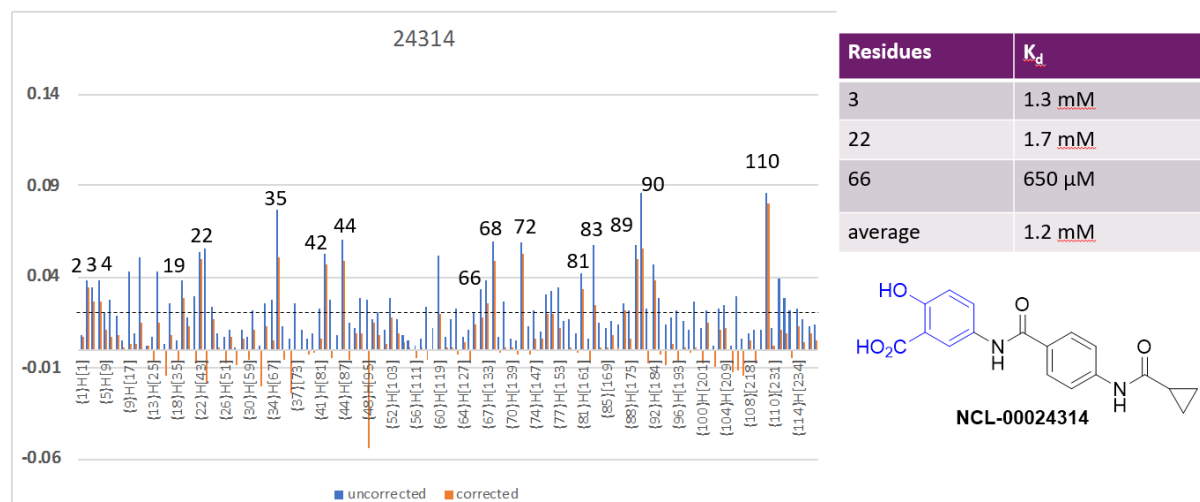


Figure 52; NMR shifted signals for the amino-acids present in the allosteric site of Sos upon addition of compound (15)

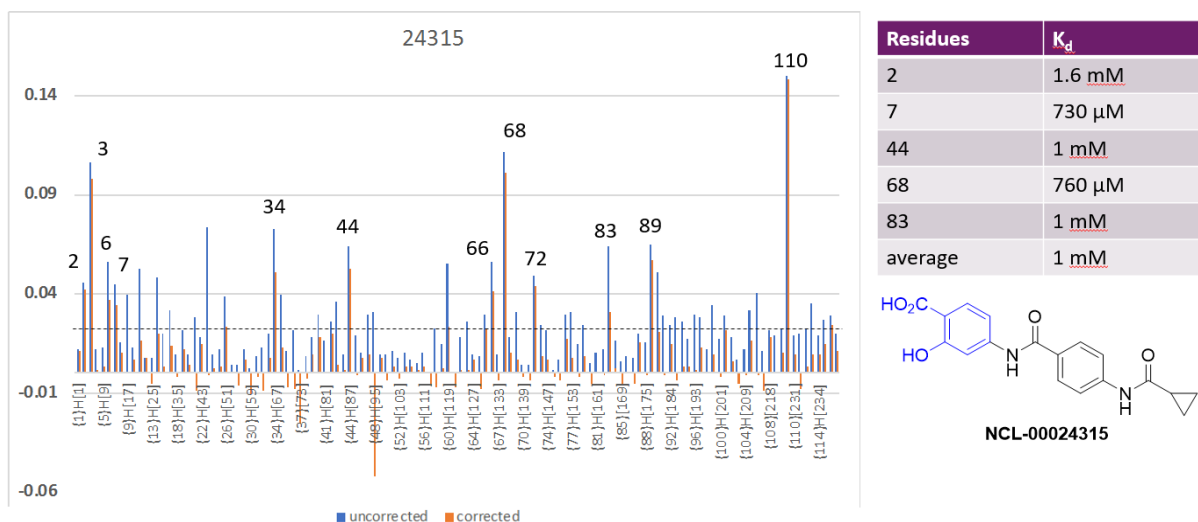


Figure 53; NMR shifted signals for the amino-acids present in the allosteric site of Sos upon addition of compound (17)

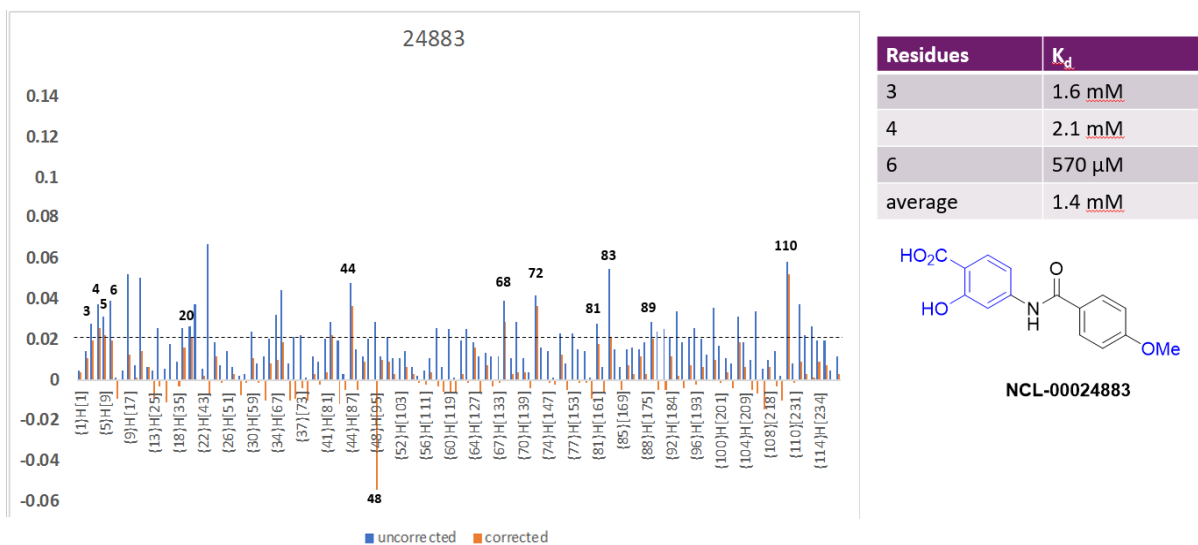


Figure 54; NMR shifted signals for the amino-acids present in the allosteric site of Sos upon addition of compound (36)

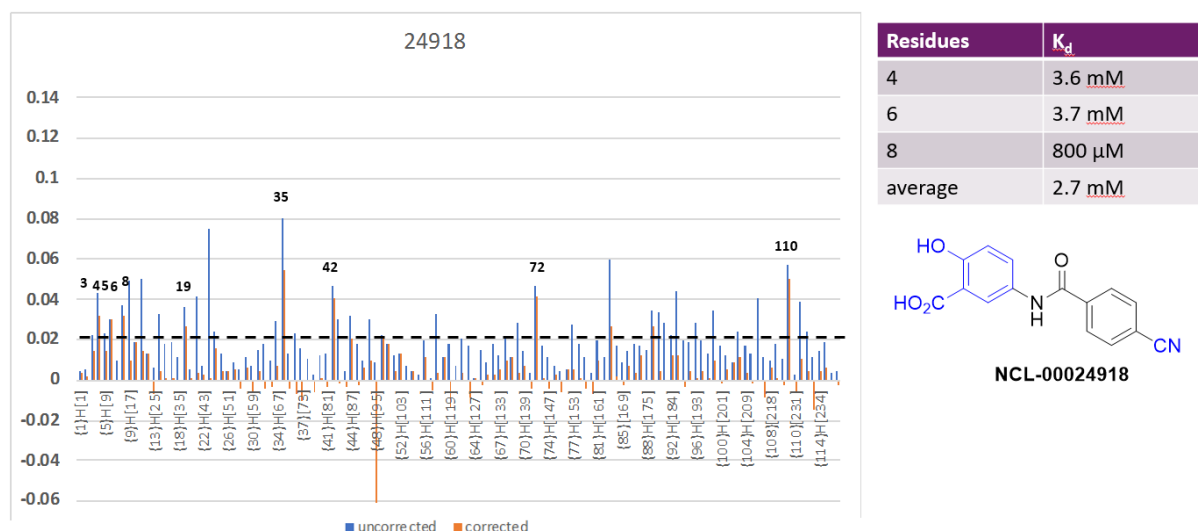


Figure 55; NMR shifted signals for the amino-acids present in the allosteric site of Sos upon addition of (37)

Regarding the affinity of our ligands to the allosteric site, with the results that we have so far only a few conclusions can be made. The cyclopropyl amide moiety is not strictly necessary for affinity, when this group was substituted by a methoxy group or a nitrile the potency slightly dropped, suggesting the introduction of a (cyclopropylmethoxy)benzene or a (cyclopropylnitrile)benzene (Figure 52 and 53).

ITC and SPR assays were also run, but the lack of solubility and the low affinity of our ligands did not give us any useful results to direct the synthesis of further analogues.

## Chapter 6: Conclusions and Future Work

The development (or choice) of a different biophysical technique with less demanding solubility conditions could give us information about the affinity of the ligands explored. For this, a lower concentration of ligand was proposed for future SPR trials.

Obtaining the crystal structure for (H-Ras:Sos<sup>cat</sup>) with soluble ligands can give us useful information for future explorations. The orientation of the ligands binding the protein, and rationalizing which substituents are necessary for binding could lead future explorations.

Also, the application of the FragLite library (see Chapter 15) by soaking the FragLites into protein crystals can give us valuable insights about the different binding pockets of our protein.

Unfortunately, without more information about the affinity of our ligands for the allosteric site we can only focus our efforts on the synthesis of new series with improved solubility.

## **Chapter 7: Introduction to inhibition of IRF4 with drug-like small molecules as a useful therapy in lymphoid cancers**

### **7.1 Transcription Factors and cancer**

#### **7.1.1 IRF family**

**Interferon Regulatory Factors (IRFs)** are proteins with the ability to block (repressor) or promote (activator) the expression of interferons (IFNs). IFNs are a class of cytokines, *i.e.* molecules used by cells to communicate with each other to trigger the immune system and eradicate pathogens. A virus-infected cell releases viral particles that can infect nearby cells, however, the infected cell can prepare neighboring cells against a potential infection by the virus by releasing IFNs. The way to prepare this response is done via motifs found in pathogens such as viral glycoproteins, lipopolysaccharides (bacterial endotoxin), bacterial flagella or viral RNA. These motifs can trigger the interferon response by binding to transmembrane receptors of cells, which can initiate several signaling pathways such as MAP kinases, PI3-K or Jak-STAT.<sup>133</sup> The final step of the signaling cascade is performed by a Transcription Factor (in the case of IFN an Interferon Regulatory Factor). IRFs can bind specific sequences of DNA and control the rate of transcription from DNA to messenger RNA. Transcription Factors can perform this function alone or as part of a complex (Figure 56).<sup>134</sup>

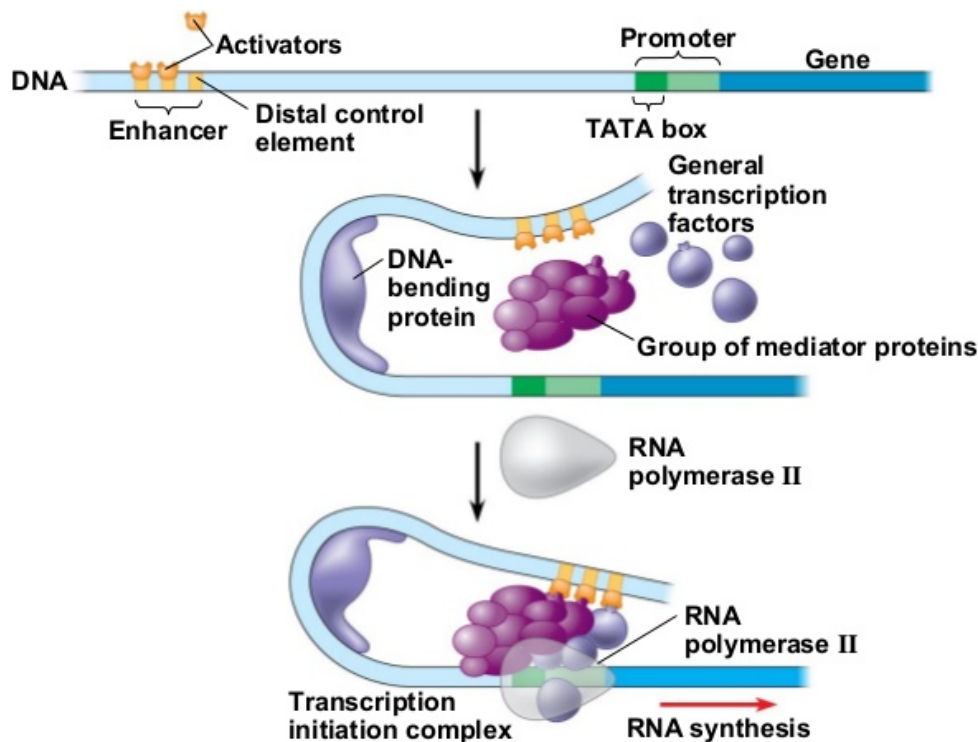


Figure 56; Transcription factor promoting RNA synthesis

Since the discovery of the first IRF transcription factor, the IRF family of proteins has grown remarkably in number and also in the diversity of functions related to them. Despite being originally defined for their relationship with viruses (notoriously Kaposi's sarcoma) and IFNs induced activation of gene expression, it is now established that the IRF proteins have a great impact in our immune responsiveness, cell growth regulation and hematopoietic development.

The IRFs can be classified into four groups, those that activate (IRF-1 and ISGF3 $\gamma$ ), those that repress (IRF-2, ICSBP), those that both activate and repress (IRF-4) and those that may activate or repress (IRF-3) transcription of target genes. All members of this family share significant homology in the N-terminus part (115 aminoacids) which comprises the DNA binding domain (DBD); through this DBD the IRFs bind similar DNA motifs termed Interferon Stimulated Response Element (ISRE; found in most IFN-inducible gene promoters), Interferon Consensus Sequence (ICS) or Interferon Regulatory Factor Element (IRF-E).

Studies characterizing IRF-expressing cell lines and IRF knockout mice reveal that each member of the IRF family performs different roles in cellular processes such as pathogen

response and cytokine signalling but can be also involved in cell growth regulation and hematopoietic differentiation. Its overexpression has been also related to cancer malignancies.

### 7.1.2 IRF-4 and cancer

Interferon Regulatory Factor 4 (Figure 57), also known as MUM, was found to be associated with sensitivity of skin to sun exposure, it has a strong presence in blue eyed, brown haired and freckled people and it is implicated in several types of leukaemia.

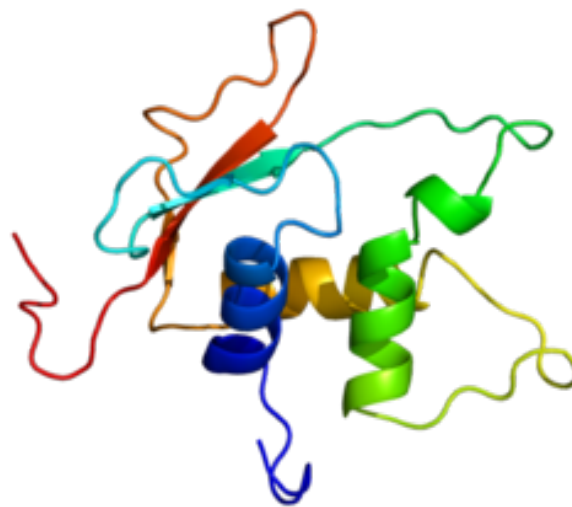


Figure 57; IRF4 Ribbon-3D structure, PDB ID: 2DLL

IRF4 is also closely related to several immune malignancies:

- Recent studies have shown that B-Cell chronic lymphocytic leukaemia (B-CLL) is the most common form of leukaemia in the Western world.<sup>135</sup> Absence of IRF4 expression is the highest relative risk among the factors analysed in determining the probability of death in patients with B-CLL. Patients without IRF4 expression had significantly worse overall survival than did those with IRF4 expression.<sup>136</sup> The overexpression or lower expression of IRF4 in the different types of diffused large B cell lymphoma has been demonstrated to be of paramount importance, however the pathogenesis and its direct relation with IRF4 is still not completely defined.<sup>137</sup>
- The overexpression of IRF4 is not confirmed to be involved in the Hodgkin and Reed Sternberg (HRS) cells, however some studies have linked the viability of these cells to the expression of IRF4.<sup>138</sup>
- In B cell acute lymphoblastic leukaemia, IRF4 is known to play a tumour suppressive role.<sup>139</sup>

- In chronic myeloid leukaemia IRF4 is shown to be under-expressed, however the role of IRF4 in the disease is not characterized yet.<sup>140</sup>
- IRF4 is associated with multiple myeloma (MM) leukaemia, and it has been commonly shown to be overexpressed without genetic alterations in a majority of multiple myeloma cases. The precise mechanism for pathogenesis of multiple myeloma in presence of high levels of IRF4 is mediated by an auto-regulatory loop established between IRF4 and c-MYC during B cell activation. Genetic abnormalities of MYC upregulate its expression in myeloma, thereby augmenting IRF4 expression. In normal plasma cells, Blimp-1 represses MYC, but this control circuit is abrogated in myeloma (Figure 58).

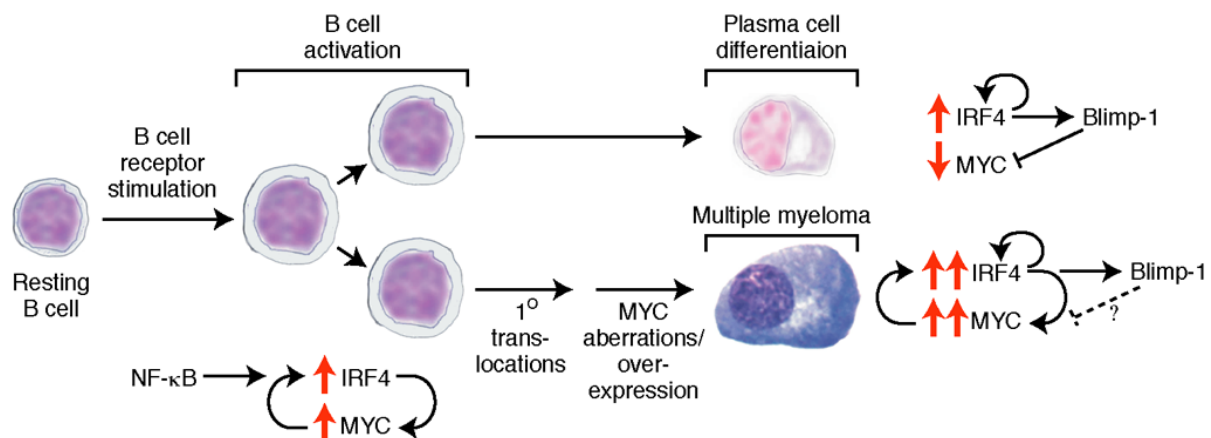


Figure 58; IRF4 and MYC autoregulatory loop during normal B cell activation

Despite functioning as a tumour suppressor gene in early B-cell development, IRF4 is a well-established oncogene in multiple myeloma, with oncogenic implications extending to certain adult lymphomas and leukaemias.<sup>141,142,143</sup> IRF4 appears to be a master-regulator of gene transcription in myeloma, influencing cellular survival, proliferation, metabolism and differentiation (Figure 59).



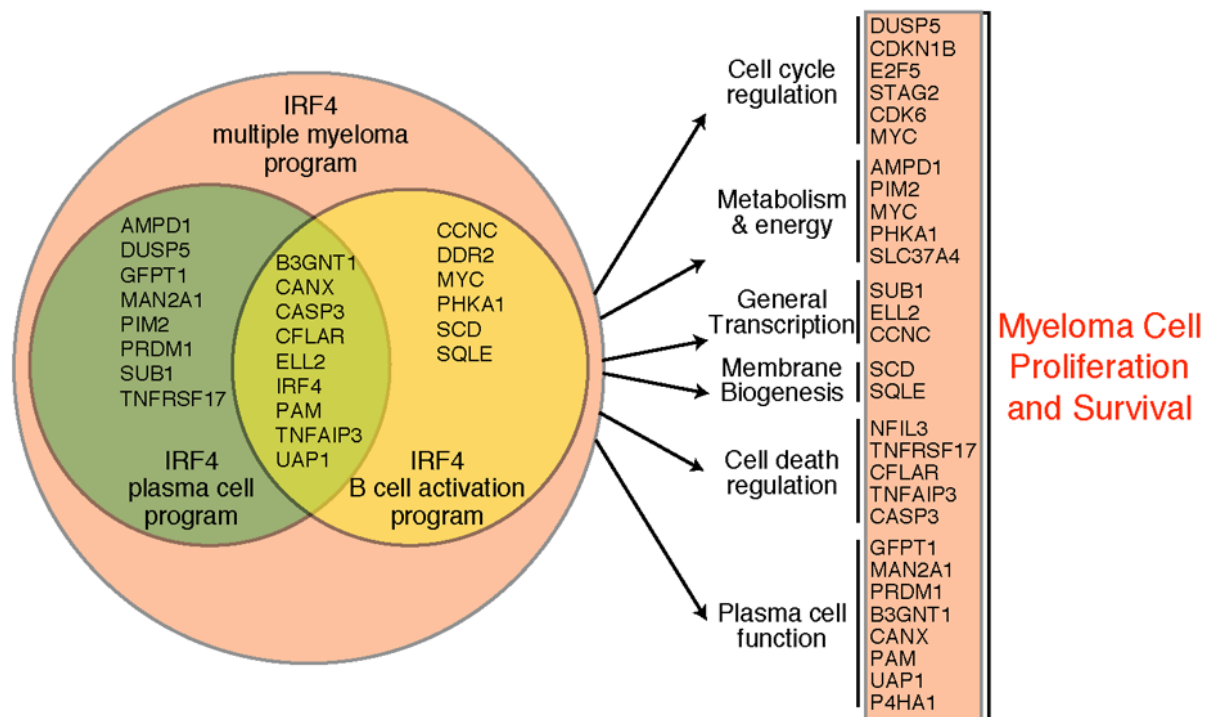


Figure 59; IRF4 Master Regulator in Multiple Myeloma

siRNA knockdown of IRF4 by only 40-60% is lethal to myeloma cells, and therefore termed “non-oncogenic addiction”. The possibility of inhibiting IRF4 is considered a highly attractive therapeutic strategy since IRF4 gene deletion results in mice that are phenotypically normal suggesting a therapeutic window.

To successfully inhibit IRF4 we need to understand how this transcription factor binds to DNA. The structure of IRF4 and its mechanism to interact with DNA have been widely studied.<sup>144</sup> It is known that the Ets (E26 transformation-specific or E-twenty-six) and IRF transcription factor families contain structurally divergent members, which have evolved to cooperatively assemble on composite Ets-IRF DNA elements to regulate gene expression in the immune system.

PU.1 is a hematopoietic Ets family transcription factor. Together PU.1 and IRF-4 cooperatively bind to composite elements found in the DNA promoters and enhancers of B-lymphoid and myeloid genes.

The interaction of PU.1 and IRF-4 involves two protein-protein interactions (Figure 60).

1. The Ets (E26) and IRF DNA binding domains

2. The phosphorylated PEST region of PU.1 and an autoinhibitory domain in IRF-4.

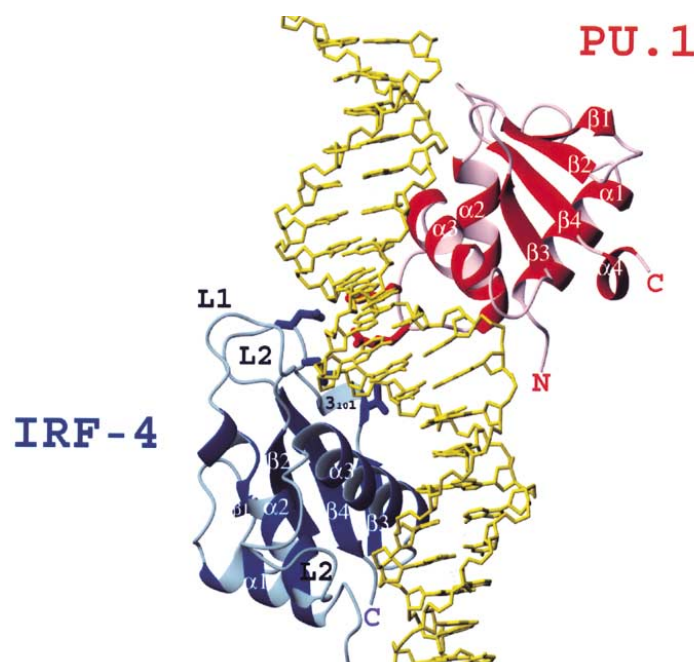


Figure 60; X-Ray Structure of ternary complex<sup>145</sup>; The PU.1 Ets domain (red) and IRF-4 DNA binding domain (blue) are accommodated on opposite faces of the DNA (gold).

Whereas PU.1 binds independently to its recognition site in DNA with high affinity, IRF-4 interacts weakly with template DNA in the absence of PU.1. The reduced affinity of IRF-4 for its binding site can be explained by its structurally diverged IRF-4 domain and by the presence of a C-terminal autoinhibitory domain.

It has been proposed that recruitment of IRF-4 DNA bound to PU.1 involves cooperative interaction between the Ets and IRF DNA binding domain (DBDs) as well as relief of autoinhibition, which is mediated by the interaction of the phosphorylated PEST region in PU.1 with the C-terminal regulatory domain of IRF-4.

In previous structure determinations done via co-crystallization of a ternary complex using murine PU.1, mouse IRF-4 and DNA fragments, the structure revealed that PU.1 and IRF-4 bind to opposite faces of DNA in a head-to-tail orientation. A continuous van der Waals surface extends between PU.1 and IRF-4 (Figure 61).

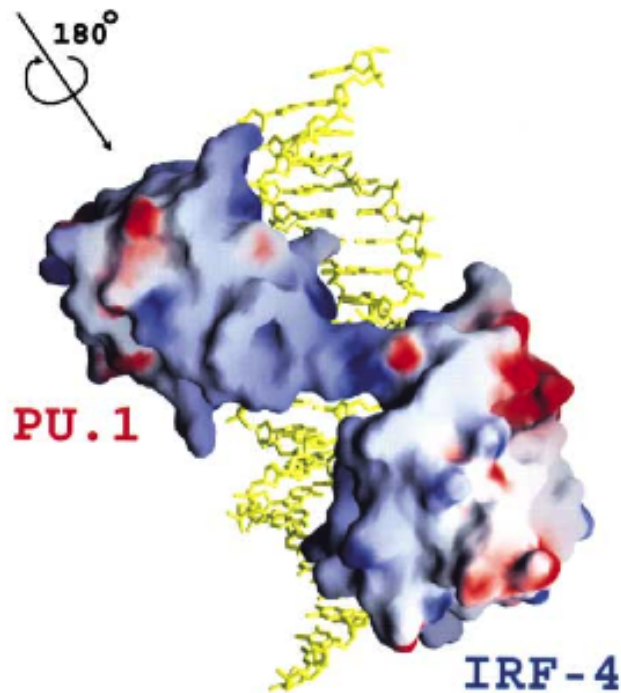


Figure 61; Computed with program GRASP<sup>146</sup>; Surface colored according to electrostatic properties, red to blue spans -10 to +10 kT in electrostatic potential.

Summarizing, IRF4 is composed by a DNA Binding Domain (N-terminal), which weakly binds to DNA, and by an Interferon Association Domain (C-terminal), which binds to PU.1. If we can interfere with the association of IRF4 with DNA, the function of the interferon regulatory factor as a transcription factor could be disrupted and potentially benefit cancer patients with IRF4 overexpression. Both association domains are relevant and potential tractability could offer an interesting therapeutic window.

## 7.2 Transcription factors as drug targets

Transcription factors could be considered as the most direct targets for treating cancer. The amount of signaling proteins affected in cancer is considerably higher than the number of oncogenic transcription factors identified, therefore an effective anti-transcription factor-drug could be useful for several upstream proteins.

Three main groups of transcription factors are known to be relevant in cancer:

1. Steroid receptors (estrogen receptors in breast cancer for example or androgen receptors in prostate cancer).
2. Resident Nuclear Proteins (usually activated by serine kinase cascades).<sup>147</sup>

3. Latent cytoplasmic factors (usually activated by receptor-ligand interaction at the cell surface).

Despite the expectation that inhibition of the interaction between DNA-binding proteins with DNA is difficult, intense screening campaigns for this purpose are being undertaken from academia and industry.<sup>148</sup> In fact, even if the specific DNA binding of a specific transcription cannot be disrupted, the interruption of functions of the transcription factor could still give us the desired pharmacological effects.

### **7.3 Current Approaches to inhibit IRF4 for multiple myeloma malignancies**

In general, only considering patients affected by MM, the 3-year survival rate decreases from 82.5% to 50.4% in patients with low and high IRF4 expression respectively.<sup>149</sup> Thus IRF4 is a target for antitumour drugs to treat multiple myeloma malignancies that has been already investigated. Research focused on the disruption of IRF4 overexpression has yielded 2 molecules with inhibitory activity and 2 different mechanism of action: lenalidomide and bortezomib.

#### **7.3.1 Lenalidomide**

Lenalidomide is an orally active compound with immunomodulatory and antiproliferative agent characteristics with several uses for haematological disorders, non-Hodgkin lymphoma and MM.<sup>150</sup> It is sometimes used in combination with the corticosteroid dexamethasone in MM patients, however a low dose of dexamethasone generally results in a superior overall survival.<sup>151</sup> These results indicated that dexamethasone adds toxicity and may limit some of the beneficial effects. Lenalidomide success relies on the inhibition of the proliferation of MM cells and immunomodulation against tumour cells. To inhibit the proliferation of MM cells lenalidomide has been proven to upregulate apoptosis via inhibition of CDK2, 4 and 6 activity which leads to retinoblastoma (RB1) hypophosphorylation and cell cycle arrest. In the case of MM patients RB1 is predominantly expressed in the phosphorylated form. Regarding the immunomodulatory actions, lenalidomide can downregulate cytokines (such as IL-6) and therefore reduce the expression of IRF4 present in cells.

Lopez-Girona, A. *et al.* (2011)<sup>152</sup> studied the association between IRF4 and MYC. It was confirmed that lenalidomide treatment decreased the amount of IRF4 and MYC proteins bound to the specific sites of IRF4 and c-MYC genes suggesting that lenalidomide interferes

with IRF4/MYC auto-regulatory loop which affects the proliferation of cells. These results support the hypothesis that IRF4 plays an important role in the inhibition of the cell cycle and proliferation of MM cells and that lenalidomide can interfere with the association with MYC rather than inhibiting the association with DNA. In the same study it was also established with a panel of MM cell lines with different sensitivity to lenalidomide that there is a relationship between IRF4 levels and lenalidomide efficacy. An overall survival analysis done on samples from 154 well-characterized MM patients confirmed a significantly shorter overall survival for patients with high IRF4 expression. Without any treatment median overall survival was 42 and 83 months in patients with high and low IRF4 expression respectively (Figure 62-A). Among patients treated with lenalidomide, there was no longer any significant difference in overall survival between patients with high or low expression of IRF4 (Figure 62-B). The importance of high IRF4 expression was even more important for patients treated with non-lenalidomide strategies (Figure 62-C).

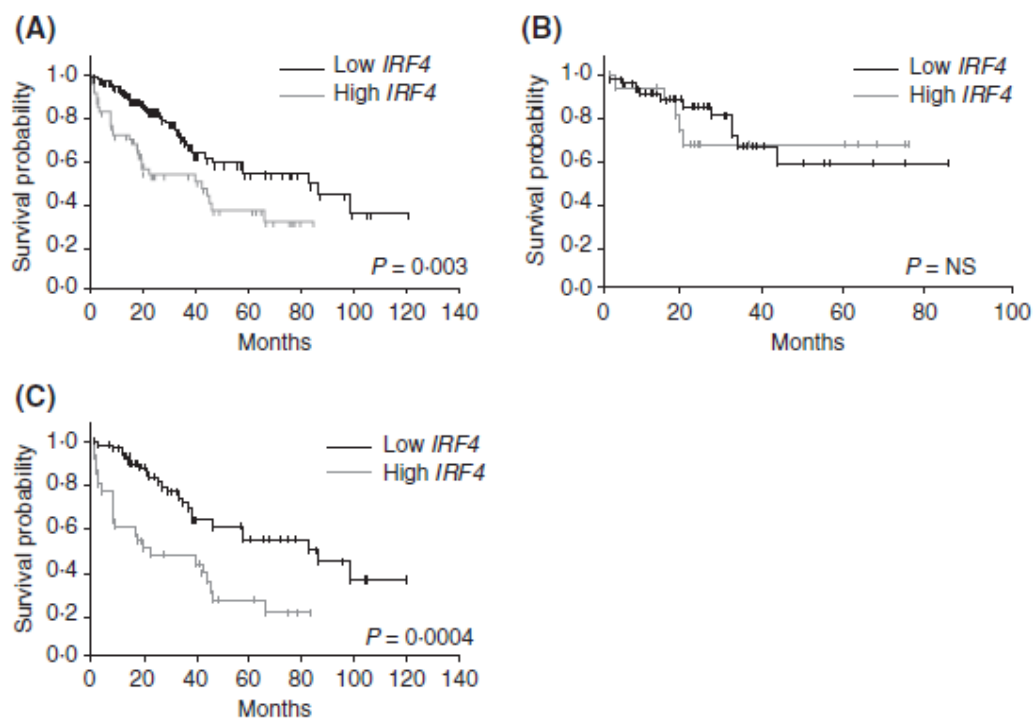


Figure 62; Treatment influence in patients with variable IRF4 expression: A) patients with no treatment, B) Patients treated with lenalidomide, C) Patients treated with other strategies

### 7.3.2 Bortezomib combined with lenalidomide<sup>153</sup>

In the case of Mantle cell lymphoma (MCL), which leads to aberrant overexpression of cyclin D1, conventional chemotherapy induces remission but unfortunately relapses within a few years leading to an overall median survival of 5-7 years.

Bortezomib is a first in class proteasome inhibitor of 26S proteasome and induces response rates of 30% to 50% in patients with relapsed disease independently of their sensitivity to prior treatment. However, more than half of all MCL patients are resistant or develop resistance to bortezomib. The mechanism for resistance is not yet understood but an increased expression of IRF4, among other things, has been pointed to by *in vitro* approaches.

In a recent attempt to establish if lenalidomide could reverse MCL resistance to bortezomib *in vivo*, mice inoculated with REC-1 cells (MCL cell line that develops resistance) were assigned to four different treatment arms (i/v administered):

1. Bortezomib (0.15 mg/kg) twice per week.
2. Lenalidomide (50 mg/kg) daily.
3. A combination of both agents.
4. Equal volume of vehicle (saline).

In the case of administration of bortezomib the size of the tumour was not affected. When administration of lenalidomide was done a 45% of reduction of the tumour size was observed when compared to the vehicle and finally the reduction of the tumour size was enhanced with the addition of bortezomib to lenalidomide, reaching a 37% and 67% tumour decrease when compared with lenalidomide and vehicle arms respectively (Figure 63).

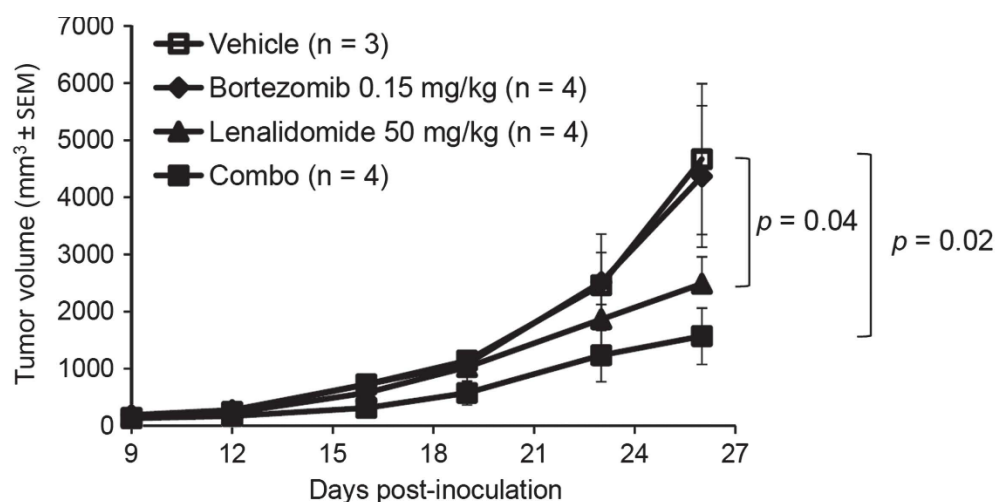


Figure 63; Mice were inoculated with REC-1 cells as previously described and began treatment at day 9 post-inoculation with 50 mg/kg lenalidomide, 0.15 mg/kg bortezomib, both agents or equal volume of vehicle. Lenalidomide was administrated 5 days a week and bortezomib twice a week for 18 days

Further histological analysis of the tumour confirmed a stronger reduction in the mitotic index by the combination therapy (Figure 64). Left panel, intra tumour glucose uptake was evaluated in representative mice injected intravenously with an IR800-labeled 2-deoxy glucose probe 24 h prior killing, and visualized with an Odyssey infra-red scanner (Li-Cor, Lincoln, NE, USA). Right panel, relative fluorescence quantification by means of the Image Studio software (Li-Cor) shows markedly reduced glucose uptake in tumour masses from mice receiving lenalidomide or the lenalidomide–bortezomib combination, when compared with either vehicle- or bortezomib-treated animals.

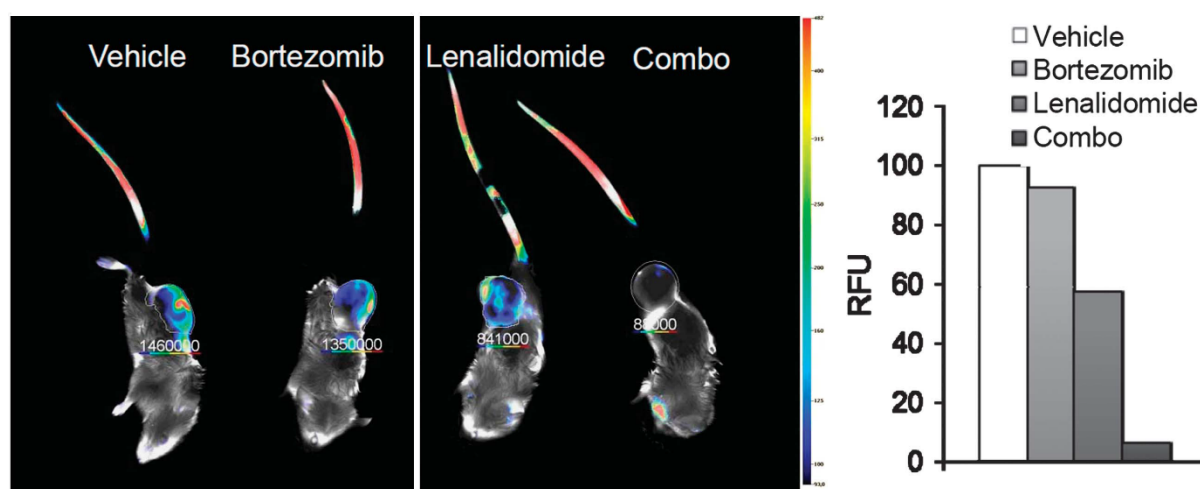


Figure 64; Infrared Scanner showing tumour volumes (blue to red: increase in glucose uptake)

After analyzing the *in vivo* studies previously described with spontaneous or acquired resistance to proteasome inhibitor it was proved that drug-resistant MCL cell lines have increased tumourigenic properties *in vivo*, and that the increased aggressiveness of these tumours is directly related to the expression of a plasmacytic differentiation program, IRF4 being the major regulator.

The evidence suggests that MYC upregulation would be present in bortezomib-resistant tumours thanks to the overexpression of IRF4 which downregulation can be achieved through administration of lenalidomide. This surprising combination of drugs offers potential opportunities for the treatment of MCL patients with resistance to bortezomib.

## Chapter 8: Project Rationale

Multiple myeloma is the second most common haematological malignancy, and responsible for 2% of all deaths from cancer. Despite novel therapies for multiple myeloma becoming available (including bortezomib and lenalidomide), the outlook remains poor. High risk patient median survival is < 2-3 years, for this reason tool compounds for IRF4 could be of great benefit to the field and would be used by many investigators worldwide.

The hypothesis of our project relies on the inhibition of interferon regulatory 4 (IRF4) with drug-like small-molecules as a possible therapy for multiple myeloma. IRF4 inhibitors can be discovered by screening and optimization through structure guided medicinal chemistry.

A small-molecule inhibitor of IRF4 would be expected to have a significant impact in the treatment of several cancers including myeloma and lymphomas, immunosecretory disorders and autoimmune disorders. It would also provide clinical proof of concept for the targeting of transcription factors – a previously unexploited approach.

A crystal structure for the C-terminal “IAD” (interferon association domain) of murine IRF4 has been previously reported (Figure 65). In the same article, small angle X-ray scattering (SAXS) data for full-length IRF4<sub>20-132</sub> are also presented and used to derive the full-length protein (in the absence of DNA).

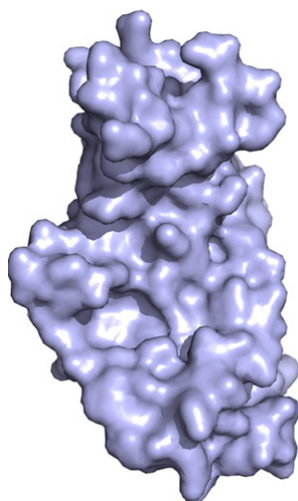


Figure 65; PDB 5BVI<sup>154</sup>



With this information, reproducible crystals of a mutant of the IRF4 DNA-binding domain (IRF4<sub>20-132</sub> E45A E46A K47A C99S) have been obtained in Newcastle University (E = Glutamate, K = Lysine, C = cysteine) by Dr. Julie Tucker (NICR).

The crystals tend to grow as clusters of plates, some of which are comprised of multiple layers (Figure 66).

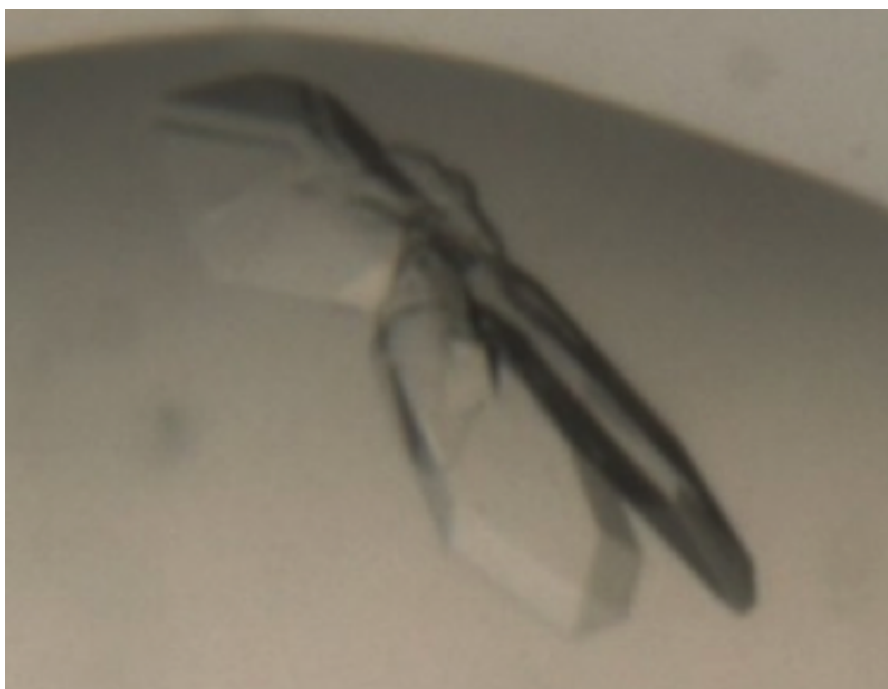


Figure 66; Crystals of IRF4<sub>20-132</sub> E45A E46A K47A C99S grown from 0.2 M NaCl, 20 % PEG3350, 0.1 M HEPES, pH 7.5

Unfortunately, the diffraction quality of these crystals is inconsistent, with some showing diffraction to better than 2.4Å on an in-house rotating anode X-ray source, whilst others show only ~3 Å diffraction at a synchrotron source (DIAMOND Light Source, Oxford). The crystals belong to at least two different space groups ( $C222_1$  and  $P2_12_12_1$ ), even though they look very similar to the eye under the microscope. The two space groups occur in the same drop under the same conditions. A subset of the crystals shows a form of twinning called “order-disorder” twinning.

In order to obtain crystals suitable for carrying out a fragment screen several conditions are mandatory:

- 1- Crystals must diffract consistently and ideally to high resolution better than 2.5 Å
- 2- Must be reproducible
- 3- Have binding sites/surfaces accessible for fragment binding
- 4- Be tolerant to high concentrations of co-solvent (DMSO) and compound
- 5- Be stable in a buffer compatible with compound solubility and binding

Due to these restrictions alternative crystallisation conditions were tried, and also some alternative protein variants were synthesised and tested. Finally, mutation of cysteine 99 to serine, to generate the quadruple mutant IRF4 (20-132 [E45A E46A K47A C99S]), led to the development of a highly reproducible, well diffracting crystal system suitable for screening (Figure 67).

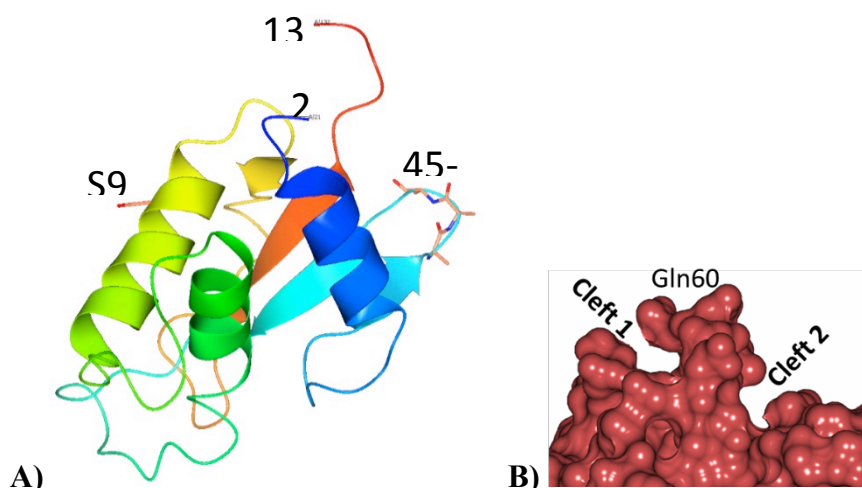


Figure 67; X-ray structure (IRF4); **A)** Ribbon diagram; **B)** DBD surface showing possible binding sites

The obtention of these crystal systems can be used for several purposes, such as:

- The completion of a model, analyzing accessible surfaces and crystal contacts
- Comparing our structure with the NMR structure of the ternary complex of IRF4 with PU.1 and DNA and DNA binding domains of other members of the IRF family
- Fully exploit the structure through pocket finding and docking
- Characterise hits from the HTS
- Run a crystallographic fragment-based screening approach to identify additional hit matter

A number of constructs spanning the IRF4 DBD variously tagged so as to facilitate the development of a variety of assays (including FP, SPR and NMR) as well as crystallisation have been generated at Newcastle University. Protein observed NMR (in collaboration with the University of York), isothermal titration calorimetry (ITC) and differential scanning fluorimetry (DSF) methods have been used to measure DNA binding to the recombinant human IRF4<sub>20-132</sub>.

### 8.1 Structural Biology and Hit Discovery

Initially a fluorescence polarization assay using fluorescently labeled DNA oligomers to characterize a variety of combinations of DNA sequences and IRF4 fragments was developed (Figure 68).

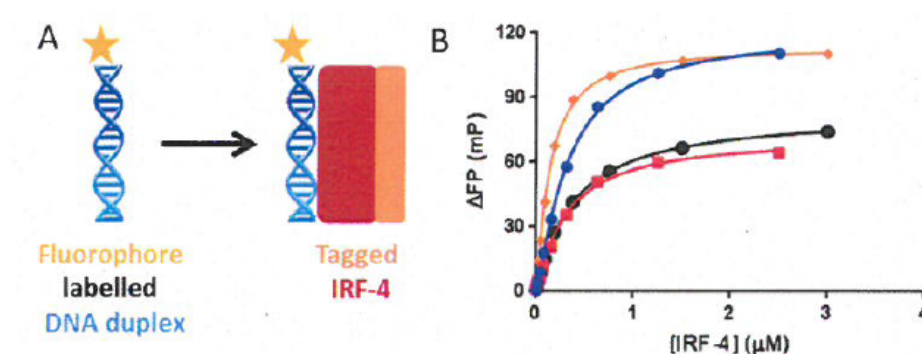


Figure 68; Fluorescence Polarization Assay; **A**) Tagging of fluorophore labelled DNA duplex with Tagged IRF4; **B**) Four different candidate molecules showing a decreased detection of fluorescence due to binding with previously IRF4 tagged-fluorophore labelled DNA for our 4 hits

A high throughput screening was carried out at CRT-DL (Cristina Alli, Fabrice Turlais, Suzanne Grooby, Cambridge) using IRF4<sub>20-132</sub> with a C-terminal hexahistidine tag and a TAMRA-labelled 12mer DNA duplex in a fluorescence polarisation (FP) assay format. The assay had been developed at NICR (Judith Reeks, Mathew Martin) based on similar assays and validated using unlabelled DNA duplex to compete off labelled DNA duplexes. Cysteine 99 within the IRF4 DNA binding domain lies close to the interface with DNA. Covalent derivatisation of Cys99 with bulky thiol-reactive compounds (S399272, 024957) was shown to abolish the signal in the FP assay, suggesting steric hindrance of DNA binding, and further validating the assay. A Cys99Ser mutant was still able to bind DNA duplex.

The current FP assay shows a lower than ideal signal-to-noise ratio with the Maximum (Protein + DNA duplex) and Minimum (DNA duplex alone) signals differing by ~1.25 fold. Further optimisation of assay parameters with the aim of improving the signal-to-noise ratio to a value  $\geq 2.5$  were carried out.

This work included testing:

- Different combinations of IRF4 construct (*e.g.* GST-tagged IRF4 instead of His-tagged IRF4) with DNA duplex (*e.g.* 21-mer instead of 12-mer).
- Alternative fluorophores (*e.g.* Fluorescein instead of TAMRA), and possibly alternative fluorophore attachment points (currently 3' to ensure close spatial proximity of fluorophore and IRF4 binding site on the duplex (Figure 69)).
- A range of buffers (*e.g.* Tris instead of HEPES), [NaCl], pH.
- The effect of additives (*e.g.* detergent, glycerol).
- The effect of temperature (*e.g.* 4°C instead of room temperature).

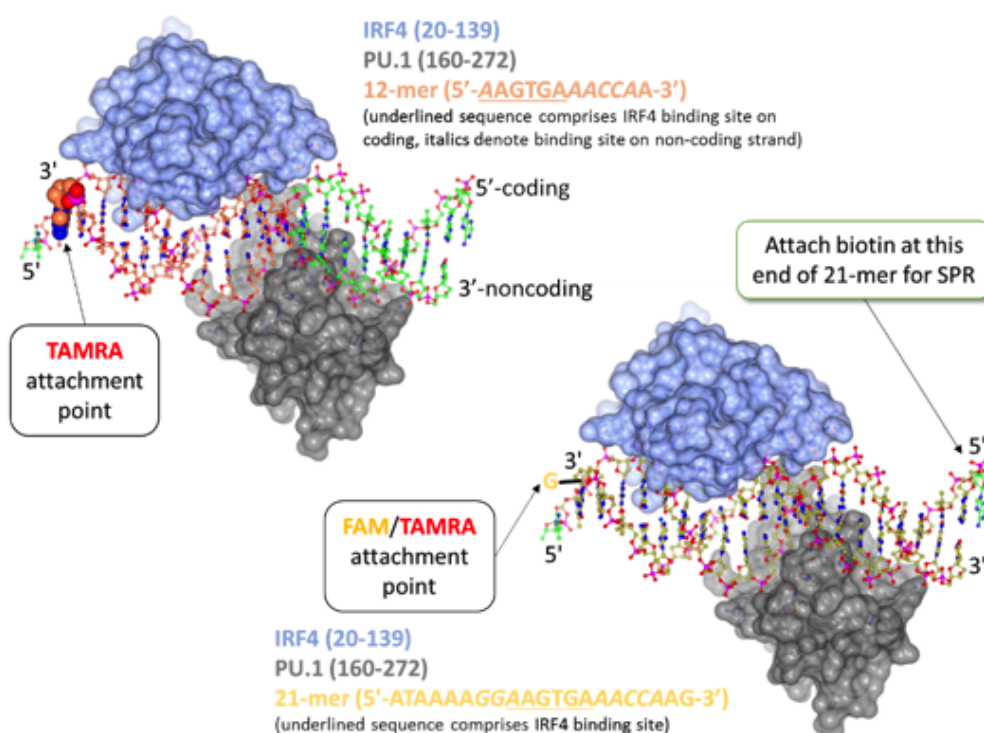


Figure 69; Design of fluorophore and biotinylated DNA duplexes for the IRF4 FP and SPR assays

Unfortunately, the FP assay could not be improved and the initial HTS hits were used for future exploration.

Four hits were selected from the initial output ((**51**), (**52**), (**53**) and (**54**)) (Figure 70), re-purchased and subjected to HPLC purification and/or re-synthesized in house.

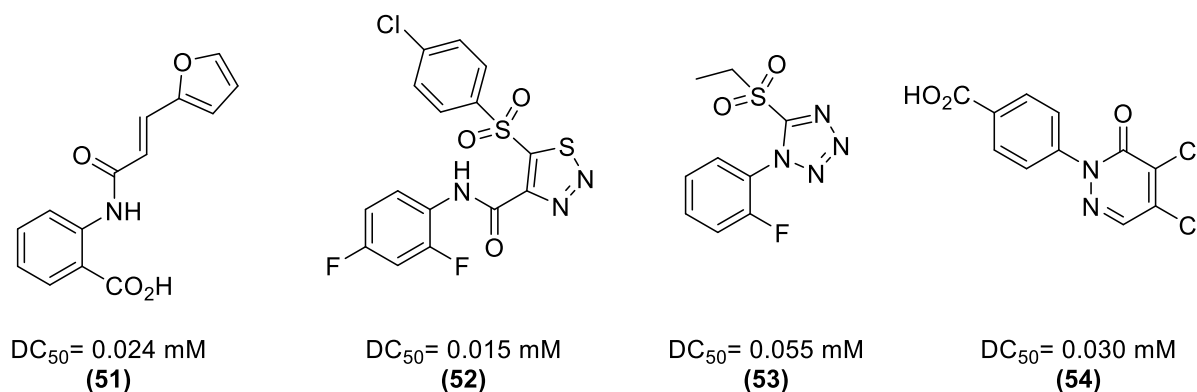


Figure 70; Hits selected after screening

The four hits were characterised in the FP assay at NICR, and shown to increase the FP signal, suggesting that they displace DNA from IRF4. Different attempts with different techniques, explained in the following pages, were performed to confirm the hits. At the same time an SAR exploration of the parent compounds (**51**) and (**54**) was undertaken.

Herein we present a summary of further biophysical assays performed to confirm the hits:

- Isothermal titration calorimetry (ITC). It is typically used as a secondary screening technique in HTS. ITC is particularly useful as it gives not only the binding affinity, but also the thermodynamics of the binding. This thermodynamic characterization allows for further optimization of compounds (Figure 71).

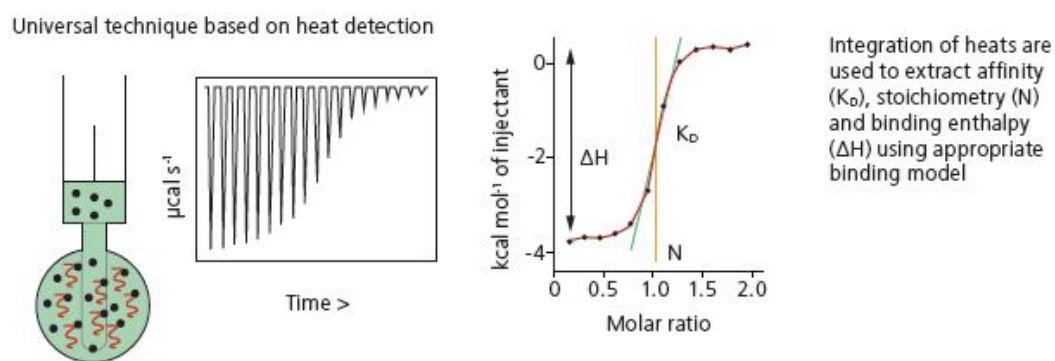
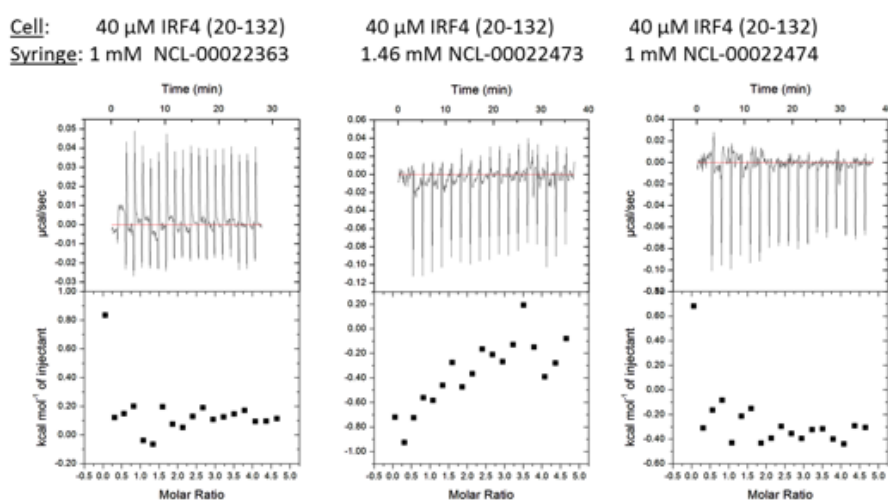


Figure 71; Basics of ITC experiments

ITC was used to assess the direct binding of the four hits to the IRF4 DNA binding domain (GPLGS-IRF4<sub>20-132</sub>-LERPHRD). Poor solubility of the compounds in aqueous buffer

supplemented with 2% DMSO limited the maximum concentration that could be achieved in the titrations. The solubility of compound (**52**) was so low that a reverse titration was performed with a lower concentration of compound in the sample cell and a higher concentration of protein in the syringe. Data for this compound were not measured at 10°C as the compound solubility was further reduced at this low temperature. The data obtained showed no evidence for the binding of NCL-00022363 (Compound (**51**)), NCL-00022472 (Compound (**52**)), or NCL-00022474 (Compound (**53**)) under the conditions used, whilst the data for NCL-00022473 (Compound (**54**)) was inconclusive (Figure 72).

Buffer system: 25mM HEPES pH7.4, 100mM NaCl, 1mM DTT, 2% DMSO  
Temperature: **10°C**



Buffer system: 25mM HEPES pH7.4, 100mM NaCl, 1mM DTT, 2% DMSO  
Temperature: **25°C**

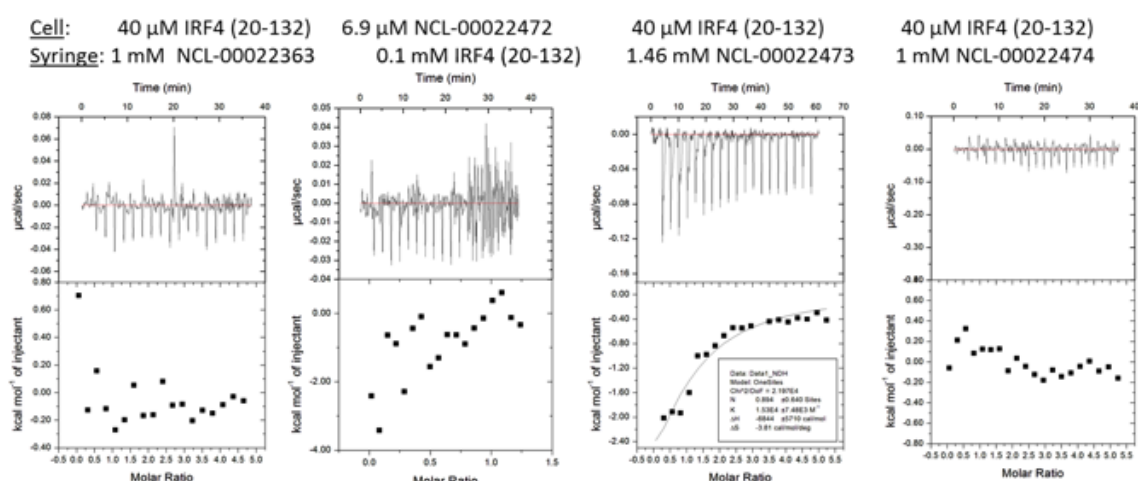


Figure 72; Data were measured using an iTC200 instrument (GE-Healthcare, located in ICaMB, Catherine Cookson building) with the indicated concentrations of protein and compound

- MicroScale thermophoresis (MST). Is based on the directed movement of particles in a microscopic temperature gradient. Any change of the hydration shell of biomolecules due to changes in their structure/conformation results in a relative change of the movement along the temperature gradient and is used to determine binding affinities. MST allows measurement of interactions directly in solution without the need of immobilization to a surface (immobilization-free technology). Preliminary investigation of the behaviour of 3'-fluorescein labelled 21-mer DNA duplex was carried out using the microscale thermophoresis instrument (NanoTemper) available in the CBCB (Baddiley-Clark Building). A sample of DNA duplex in FP assay buffer was measured. Whilst demonstrating that a good signal could be obtained, the data showed that aggregation of the DNA was likely occurring under the conditions used. Use of a detergent in the assay buffer was suggested by the MST technical support (Astrid Sitte) who helped with the measurements, but no improvement was observed.
- Nuclear Magnetic Resonance Spectroscopy. Measurements have been carried out on the 700 MHz NMR instruments (Bruker) in York (Pedro Aguiar, Jennifer Potts) and in Newcastle (Corinne Wills, William McFarlane). A number of different techniques have been used to assess direct binding of compounds to the IRF4 DNA binding domain. In York, two ligand observed methods were employed; saturation transfer difference (STD) and WaterLOGSY. Unfortunately, due to the small size of the IRF4 DNA binding domain (13 to 15 kDa depending on the protein batch used), the nuclear Overhauser transfer effects (nOe) would not be expected to be strong enough to provide a good signal in these experiments at the protein concentrations employed (up to 10  $\mu$ M). Also, critical control experiments for the WaterLOGSY experiments were not carried out, such that aggregation of the compound under the experimental conditions cannot be ruled out. Although these experiments have not so far provided evidence for direct binding of the compounds to IRF4, neither have they completely ruled this out.

In Newcastle, preparatory work to assess direct binding of compound by monitoring the chemical shift of the methyl protons in the IRF4 protein  $^1\text{H}$  spectrum was performed. An initial DMSO titration experiment shows that the IRF4 DNA binding



domain can tolerate at least 10 % DMSO without loss of signal (Figure 71). Binding of compound to the protein should impact directly on the environment of methyl groups in amino acids within the compound binding site (*i.e.* leucine, methionine, threonine, alanine, valine, isoleucine), thus causing a chemical shift. Alternatively, binding may cause a change in the overall protein conformation, which will alter the environment of methyl groups distal to the binding site. Monitoring the methyl groups is therefore a highly sensitive method of detecting binding. The DC<sub>50</sub> values for compound (**51**) and (**54**) as measured in the FP assay at NICR are in the 30-100  $\mu$ M range and unfortunately no ligand binding events were detected as we can see in the lack of chemical shift of the methyl protons (Figure 73) for compound (**54**) under any conditions.

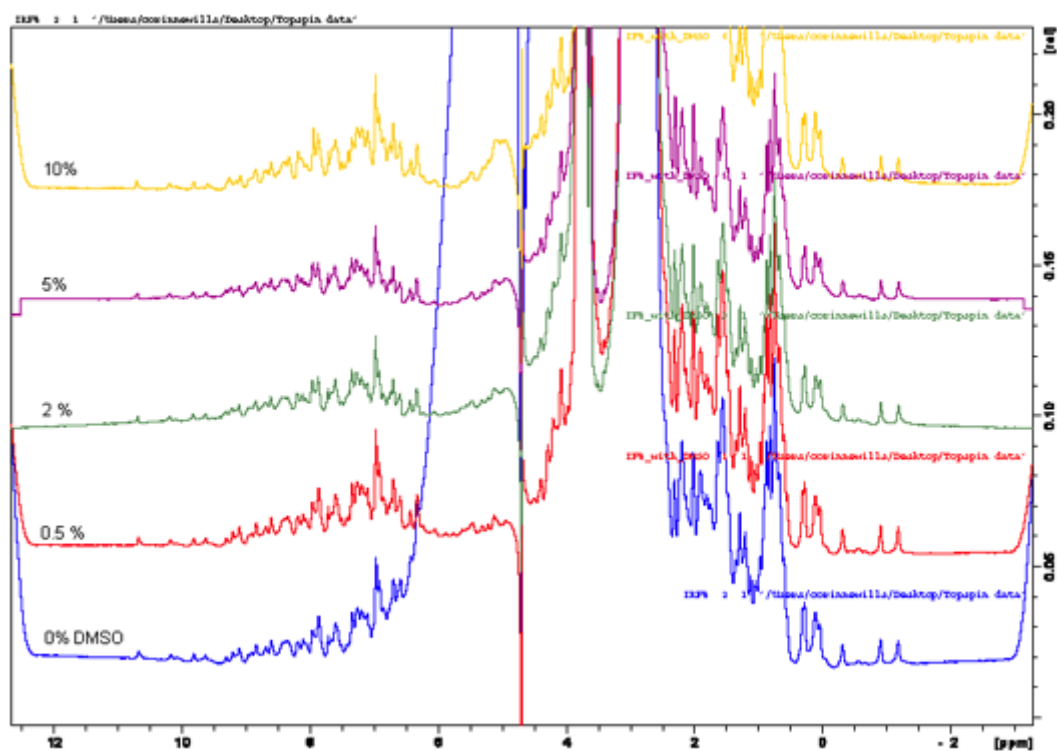


Figure 73; Titration of DMSO into IRF4 monitored by <sup>1</sup>H NMR. Methyl proton shifts are not visible at around -1.0 ppm

- Differential Scanning Fluorimetry (DSF). Changes in the protein and/or DNA duplex melting temperature ( $T_m$ ) can be used to monitor the binding of compounds. This technique has already been used in the IRF4 project, both to assess binding of DNA duplex to the DNA binding domain, and to screen a small library of fragments provided by CRT. The  $T_m$  shifts observed, including those for DNA binding, were



very small ( $< 3^{\circ}\text{C}$ ), such that the signal-to-noise is very low. The results are not reliable.

As a conclusion, none of the techniques that we tried were able to confirm any of the parent compounds as active.

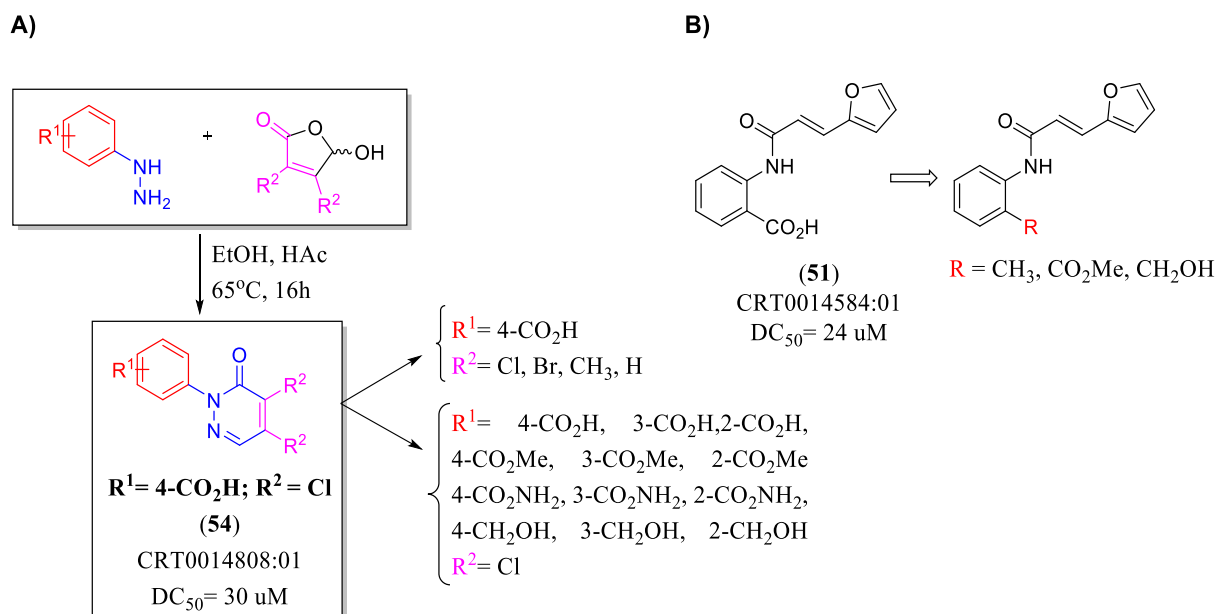
In the case of the ITC experiments we can observe some activity for compound (**54**) when the assay is tested at  $25^{\circ}\text{C}$ . Future analogues of this hit will be synthesized and further assays in the same conditions could give us some insights in the result previously obtained.

Also, in our effort to try to find compounds more suitable for the different techniques applied we decided to pursue an exploration on the scaffolds obtained during Hit Finding. The SAR conducted is presented in the following chapter.

## Chapter 9: Results and Discussion

### 9.1 Methodology

The first goal of the project was to develop robust synthetic methods for the target compounds. With the appropriate synthetic approach we were able to produce a series of analogues that can improve the potency of the compounds and establish a structure activity relationship. After the HTS several compounds were identified, of those the pyridazinone and acrylamide series were chosen for further exploration (Scheme 17).



Scheme 17; Target compounds; **A)** Pyridazinone series **B)** Acrylamide series

### 9.2 Targets synthesized

Our present effort is a follow up on HTS hits, for compound **(54)**<sup>155</sup> (Pyridazinones), and for **(51)**<sup>156</sup> (Acrylamides).

#### 9.2.1 Pyridazinones

Pyridazinone compounds have been studied as potential targets for the development of treatments of proliferative diseases such as lung cancer.

Pyridazine and its derivatives have been known for almost a century but have received not enough attention until recently due to the discovery of medicinally useful natural products.<sup>157</sup> Today, the pyridazinone nucleus and its 3-oxo derivative (2H-pyridazin-3-one) moieties have

been recognized as versatile pharmacophores. This key subunit is part of many biologically active substances with a broad range of biological and pharmaceutical activities like cardiostimulant agents<sup>158</sup> or as a treatment of neuropathic pain.<sup>159</sup>

The strategy followed initially to identify new candidates with improved properties was the exploration of parent compound (**54**). The exploration was undertaken by synthesizing analogues modifying the benzoic acid moiety (positions **R**) and the substituents on the pyridazinone ring (position **R<sub>2</sub>**) (Scheme 18).



Scheme 18; Derivatization of the structure of the parent compound

These two different series of targets were divided between me and Shaun R. Stevens. Here we present the work carried out to synthesize the **R<sub>2</sub>** analogues.

#### 9.2.1.1 Pyridazinone variations

Several derivatives have been synthesized changing the substituents of the pyridazinone ring system. A classical approach was used initially for targets (**54**) and (**55**) (Figure 74).

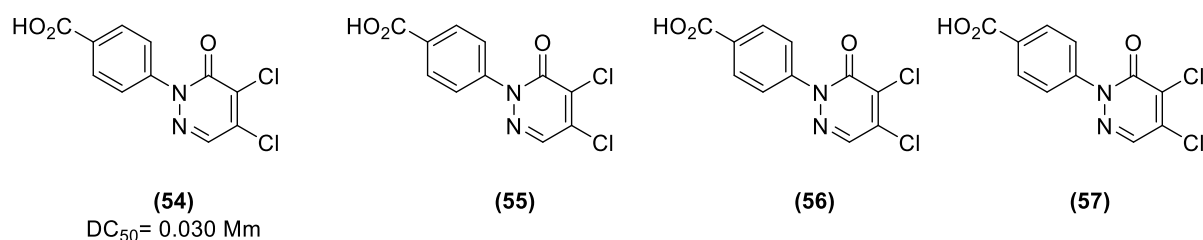
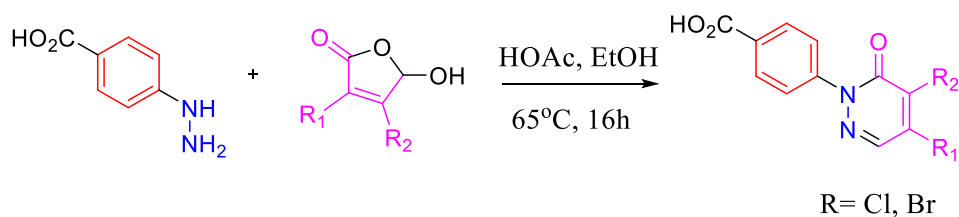


Figure 74; Pyridazinone ring derivatives

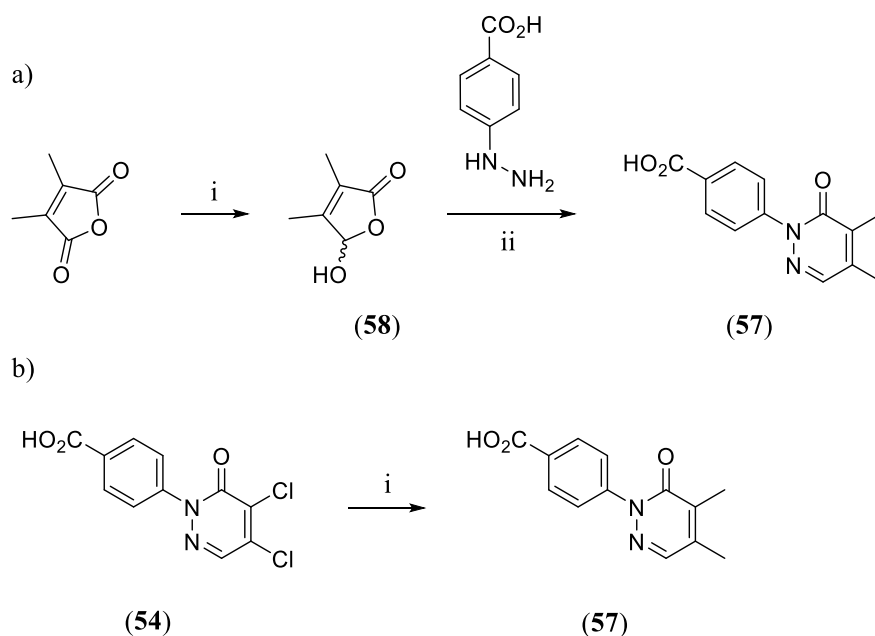
Traditionally, these compounds have been obtained by acid-catalyzed cyclocondensation reaction of N-arylhydrazine with various 1,4-difunctionalized compounds. The reactions are performed in anhydrous ethanol and with a catalytic amount of acetic acid (Scheme 19).



Scheme 19; Aryl hydrazine and 1,4-difunctional hydroxyfuranone reacted at 65°C for 16 hours

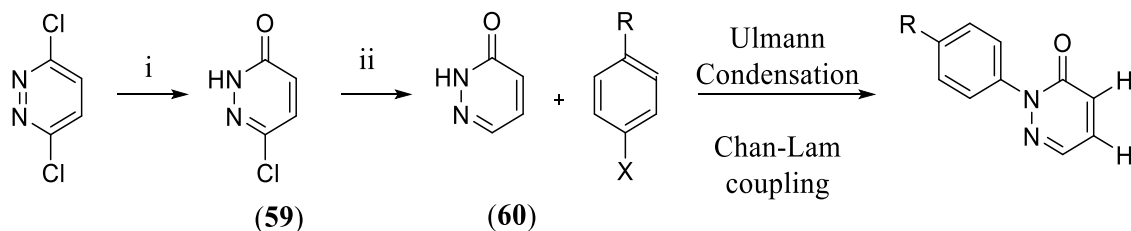
For target (**57**) two different approaches were carried and both produced the desired target (Scheme 20). In the first approach (Scheme 20-a) to synthesize the compound of interest an adaptation of the classic approach was used. 2,3-Dimethylmaleic anhydride was treated with lithium tri-*tert*-butoxyaluminium hydride in THF-Et<sub>2</sub>O at room temperature for 6 hours to afford our 1,4-difunctionalized compound (**58**). Subsequent treatment with 4-hydrazinobenzoic acid did not give the desired target (**57**).

In the second approach and as a more convergent synthetic pathway, compound (**54**) (Scheme 20-b) was used as starting material, after treatment with aluminium trichloride and palladium *tetrakis* in toluene at reflux temperature (100°C) the desired target (compound (**57**), Scheme 20-b) was obtained. In this case the yield was superior and for our interests the synthetic route is more convergent.



Scheme 20; **a**) i) Li(<sup>t</sup>BuO)<sub>3</sub>Al-H, r.t.; THF; ii) aryl hydrazine, EtOH-H<sub>2</sub>O, HAc; 60°C; **b**) i) AlMe<sub>3</sub>, Pd(PPh<sub>3</sub>)<sub>4</sub>, toluene, 100°C

For target **(56)** (Figure 74) an initial approach was developed using an Ullmann condensation as a final step for obtaining the desired target (Scheme 21).



Scheme 21; i)  $\text{K}_2\text{CO}_3$ , EtOH, reflux; ii)  $\text{NH}_4\text{HCO}_2$ , EtOAc; iii) Ullmann condensation or Chan-Lam coupling

After a wide exploration of different catalysts, bases, solvents and aryl halides, the approach using pyridazin-3(2*H*)-one (**(60)**) was finally abandoned due to lack of reactivity of the pyridazinone (**(60)**) as a nucleophile (Table 2).

Substrate	Ligand	Catalyst	Base	Solvent	Results
		CuCl	$\text{K}_2\text{CO}_3$	DMF	1% by LCMS
		CuCl	$\text{K}_2\text{CO}_3$	DMF	0% by LCMS
		$\text{CuO}_2$	$\text{Cs}_2\text{CO}_3$	$\text{CH}_3\text{CN}$	0% by LCMS
	Salox	$\text{CuO}_2$	$\text{Cs}_2\text{CO}_3$	$\text{CH}_3\text{CN}$	0% by LCMS
	$\text{Pb}(\text{OAc})_4$	$\text{ZnCl}_2$	\	Toluene	0% by LCMS
		CuCl	$\text{K}_2\text{CO}_3$	DMF	3% by LCMS
		CuCl	$\text{K}_2\text{CO}_3$	DMF	0% by LCMS

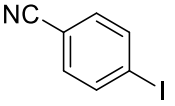
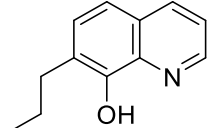
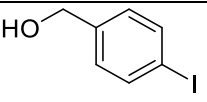
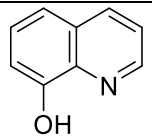
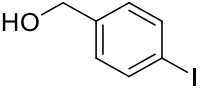
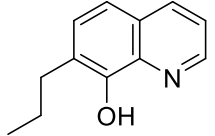
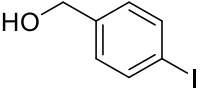
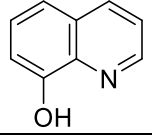
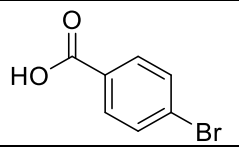
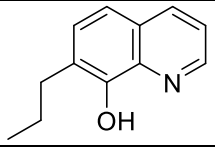
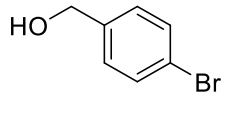
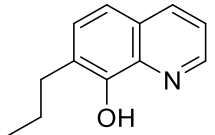
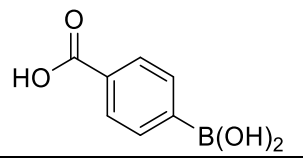
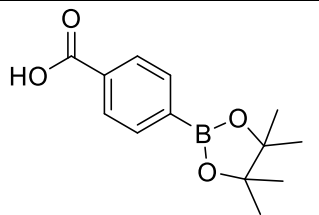
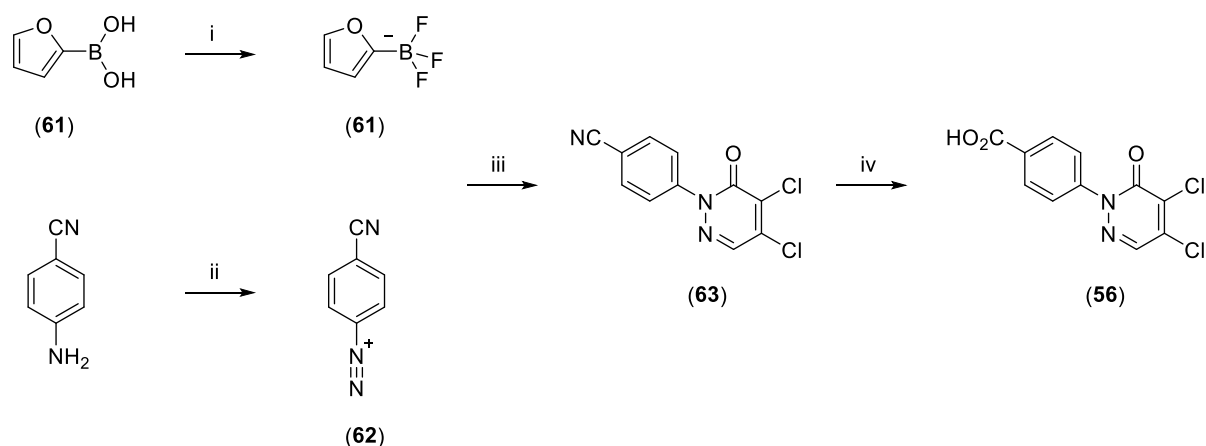
		CuCl	K <sub>2</sub> CO <sub>3</sub>	DMF	0% by LCMS
		CuCl	K <sub>2</sub> CO <sub>3</sub>	DMF	0% by LCMS
		CuCl	K <sub>2</sub> CO <sub>3</sub>	DMF	0% by LCMS
		CuO <sub>2</sub>	Cs <sub>2</sub> CO <sub>3</sub>	CH <sub>3</sub> CN	0% by LCMS
		CuCl	K <sub>2</sub> CO <sub>3</sub>	DMF	0% by LCMS
		CuCl	K <sub>2</sub> CO <sub>3</sub>	DMF	0% by LCMS
	-	Cu(OAc) <sub>2</sub>	pyridine	DMF	0% by LCMS
	-	Cu(OAc) <sub>2</sub>	pyridine	DMF	0% by LCMS

Table 2; (Ullmann condensation and Chan-Lam attempts)

A new approach was tried and found to be successful (Scheme 24). In this approach trifluoro(furan-2-yl)-borane potassium salt (**61**) was prepared from furan-2-ylboronic acid and 4-cyanobenzene diazonium fluoride (**62**) was prepared from 4-aminobenzonitrile. These two compounds were slowly mixed together in water at 5° C and after 30 minutes the reaction went to completion. After the intermediate compound 4-(6-oxopyridazin-1(6*H*)-yl)benzonitrile (**63**) was isolated and hydrolyze in acidic conditions the desired target molecule (**56**) (Scheme 22) was obtained.



Scheme 22; i)  $\text{KHF}_2$ , MeOH,  $0^\circ\text{C}$  to r.t.; ii)  $^t\text{BuONO}$ ,  $\text{BF}(\text{OEt})_3$ , DCM,  $-15^\circ\text{C}$  to r.t.; iii)  $\text{H}_2\text{O}$ ,  $5^\circ\text{C}$  to r.t.; iv)  $\text{H}_2\text{SO}_4/\text{H}_2\text{O}$  (6:4), r.t.

#### 9.2.1.2 Pyridone and Pyrimidinone variations

To obtain more information regarding the functions of the heteroatoms in the pyridazinone ring we synthesized a new series of compounds with different heterocycles. Nitrogen in position 2 of the pyridazinone ring was moved and deleted (Figure 75).

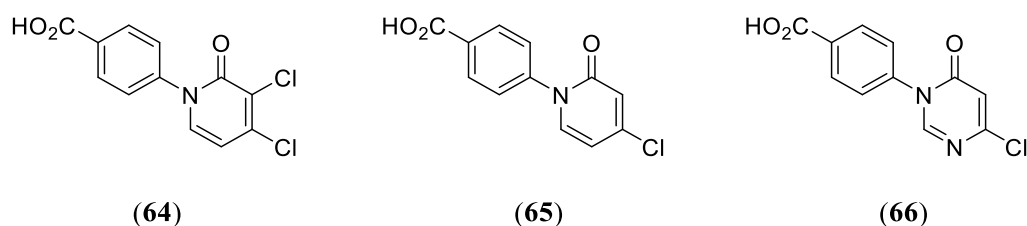
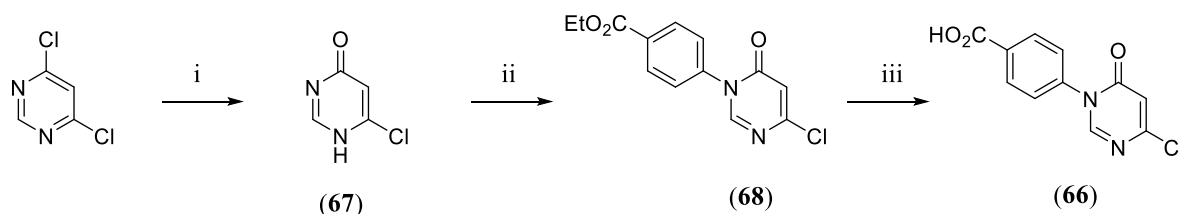


Figure 75; Pyridone and pyrimidinone ring derivatives

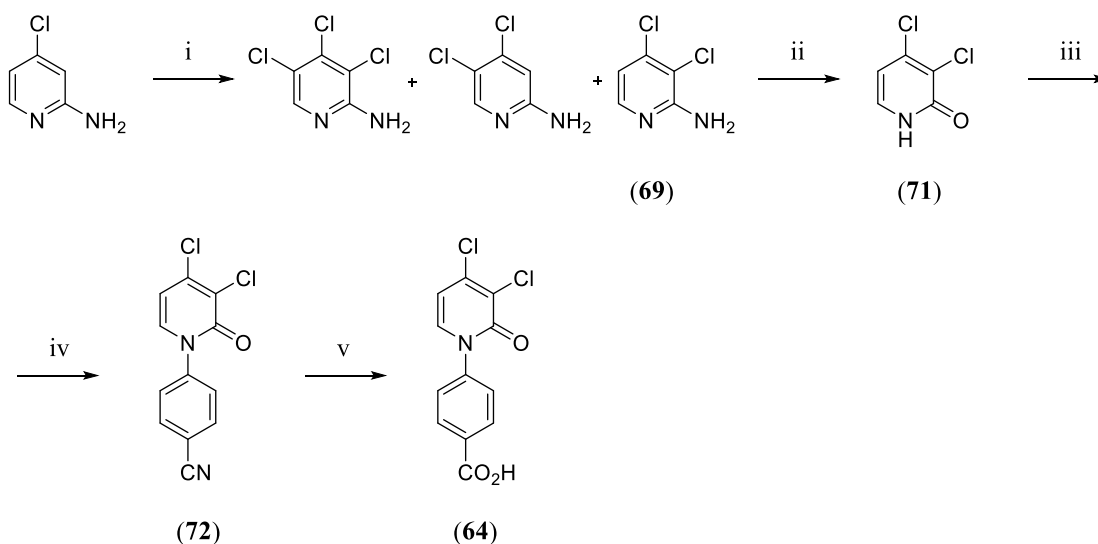
To synthesize compound (66) 4,6-dichloropyrimidine was treated with conc. HCl in dioxane. The mixture was stirred at  $70^\circ\text{C}$  for 6 hours affording compound (67) in 57% yield after purification. The pyrimidinone (67) was then reacted with 4-(ethoxy)benzene boronic acid under Chan-Lam conditions affording intermediate (68) in 62% yield. The hydrolysis of this intermediate under acidic conditions gave us the desired compound (66) in 11% yield after purification (Scheme 23-a). To synthesize compound (64), 2-amino-4-chloropyridine was treated with  $\text{H}_2\text{O}_2$  in conc. HCl at  $0^\circ\text{C}$ , affording compound (69) in 27% yield after purification. The intermediate was reacted with sodium nitrite in HCl to afford the desired pyridinone (71) in 16% yield after purification. Compound (71) was then reacted with 4-iodobenzonitrile under Ullman conditions affording the desired compound (72) in 16% yield after purification. The hydrolysis of compound (72) under acidic conditions gave us the

desired compound (**64**) in 7% yield after purification (Scheme 23-b). To make compound (**65**) 2,4-dichloropyridine was treated with sodium acetate in acetic acid to afford pyridinone (**73**) in quantitative yield. The intermediate was reacted with ethyl 4-iodobenzoate under Ullman conditions to give (**74**) in 35% yield after purification. Hydrolysis of compound (**74**) gave the desired product (**65**) in 53% yield after silica purification (Scheme 23-c).

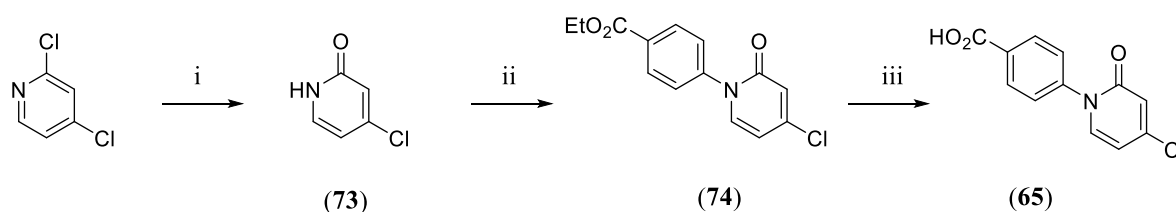
a)



b)



c)



Scheme 23; **a**) i) HCl/H<sub>2</sub>O (1:9) in Dioxane; ii) 4-(Ethoxy)benzene boronic acid, Cu(OAc)<sub>2</sub>, pyridine in DCM; iii) H<sub>2</sub>SO<sub>4</sub>/H<sub>2</sub>O (6:4); **b**) i) H<sub>2</sub>O<sub>2</sub> in HCl; ii) NaNO<sub>2</sub> in HCl; iii) 4-Iodobenzonitrile, 8-Hydroxyquinoline, CuI, K<sub>2</sub>CO<sub>3</sub> in DMSO; iv) H<sub>2</sub>SO<sub>4</sub>/H<sub>2</sub>O (6:4); **c**) i) NaOAc/ HAc; ii) Ethyl 4-iodobenzoate, K<sub>3</sub>PO<sub>4</sub>, CuI, *N,N*-Dimethylethylenediamine in dioxane; iii) LiOH.H<sub>2</sub>O in THF



Further to these analogues Shaun R. Stevens prepared compounds by growing the left-hand-side moiety and making variations in the same benzene ring. These compounds will be shown together with the FP assays at the end of this section.

### 9.2.1.3 Sulfonamide variations

After several issues with the solubility in DMSO of our compounds affecting its reactivity and having not only poor yields, but also many problems in the reliability of the biological assays, it was decided to substitute the carboxylic group in the benzylic ring for a new series of sulfonamide analogues (Figure 76).

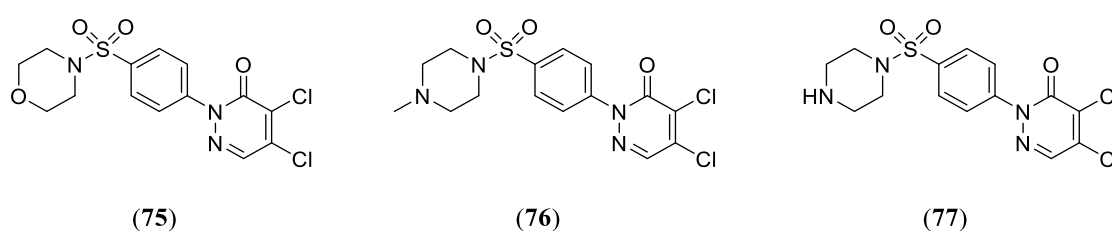
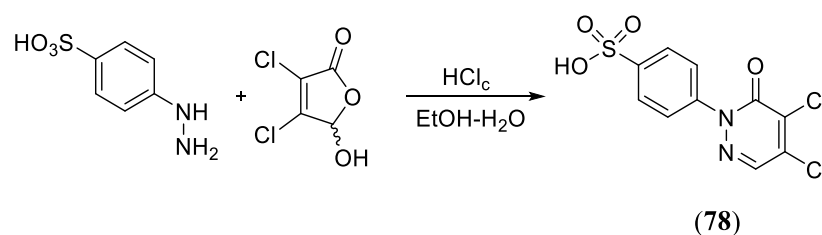


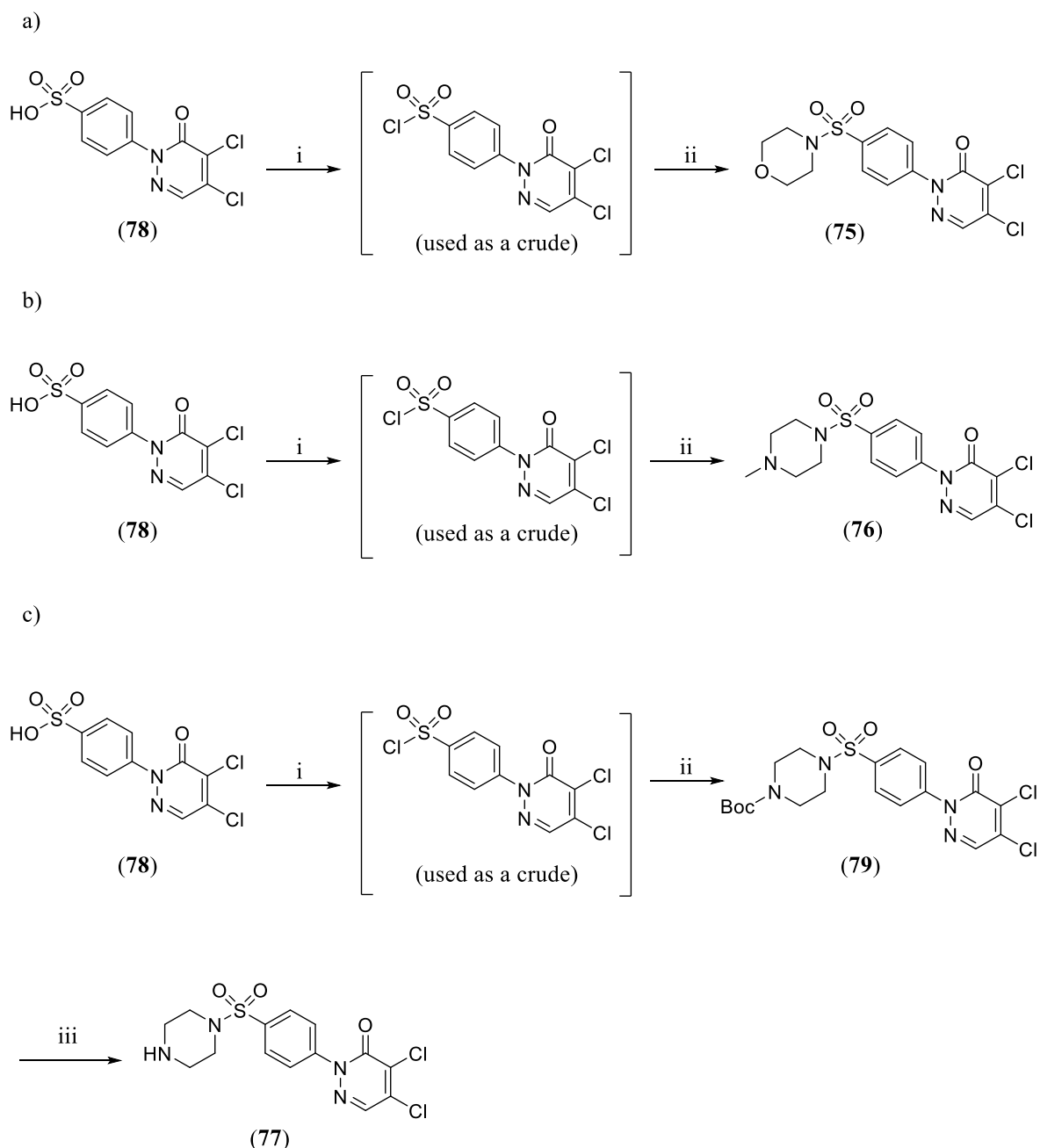
Figure 76; Sulfonamide derivatives

Our approach to synthesize these compounds was similar to that used in section 9.2.1.1 using a different hydrazine compound as starting material, this gave us access to the sulfonamide compounds. A common intermediate (**78**) was synthesized and used as precursor of the final target in the three different final targets (Scheme 24).



Scheme 24; Synthesis of common intermediate (**78**) for synthesizing sulfonamide derivatives

Once intermediate (**78**) was synthesized the derivatization for the different final targets was performed as described in the following scheme. Intermediate (**78**) was treated with thionyl chloride and the sulfonyl chloride crude intermediate was then reacted with morpholine, methyl piperazine or Boc-piperazine (Scheme 25).



Scheme 25; a) i)  $\text{SOCl}_2$  in DMF; ii) Addition of morpholine; b)  $\text{SOCl}_2$  in DMF; ii) Addition of N-methyl piperazine; c) i)  $\text{SOCl}_2$  in DMF; ii) Addition of *N*-Boc-piperazine; iii)  $\text{H}_2\text{SO}_4/\text{H}_2\text{O}$  (6:4)

### 9.2.2 Acrylamides

For the parent compound (**51**) several analogues were synthesized. Shaun R. Stevens was focused in modifying the THF moiety (left-hand side of the molecule) and I synthesized analogues substituting the carboxylic moiety for different functional groups (Figure 77).

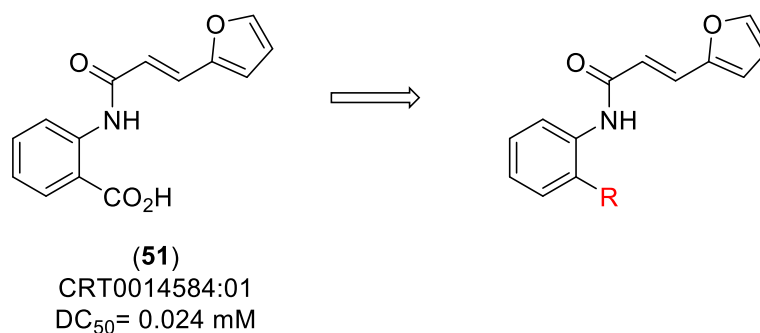
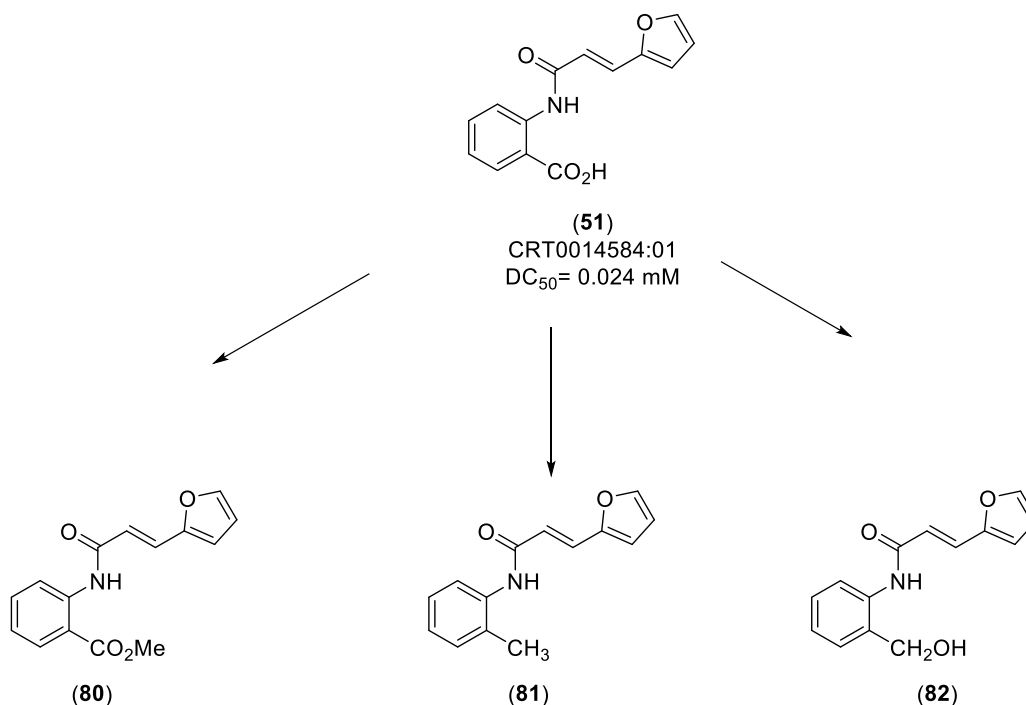


Figure 77; Derivatization on position 2 of the aniline

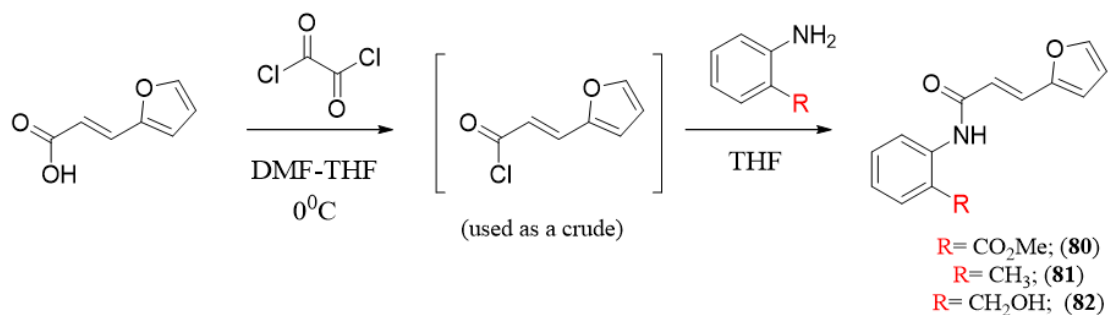
### 9.2.2.1 Acid variations

Several derivatives were synthesized with different substituents in position 2 of the initial aniline. The main purpose of these variations was to, on top of improving the potency of the scaffold, work with compounds with better solubility to obtain information in the assays. It was believed that the removal of the carboxylic group in exchange for more lipophilic groups would improve the solubility in DMSO but little or no improvement was observed (Scheme 26).



Scheme 26; Acrylamides synthesized with different substituents in position 2 of the aniline

We have used the same synthetic pathway for our desired targets in the three different analogues. (Compounds **(80)**, **(81)** and **(82)**) (Scheme 29). (*E*)-3-(furan-2-yl) acrylic acid was used as starting material in all our cases, after treatment with oxalyl chloride to give the intermediate acyl chloride, which after treatment with the appropriate aniline compounds gave the desired targets (Scheme 27).



Scheme 27; General procedure for the synthesis of acrylamides

## Chapter 10: Biophysical Assays and Conclusions after HTS and SAR focused on the DNA Binding Domain (DBD)

The hits obtained via high throughput screening (HTS) were not confirmed by the initial FP assay, which shows a lower than ideal signal-to-noise ratio with the Maximum (Protein + DNA duplex) and Minimum (DNA duplex alone) signals differing by ~1.25 fold.

Some compounds were at or, above or below, this maximum and minimum making the results inconclusive and tough to analyse. No improved activity over the parent compounds was observed.

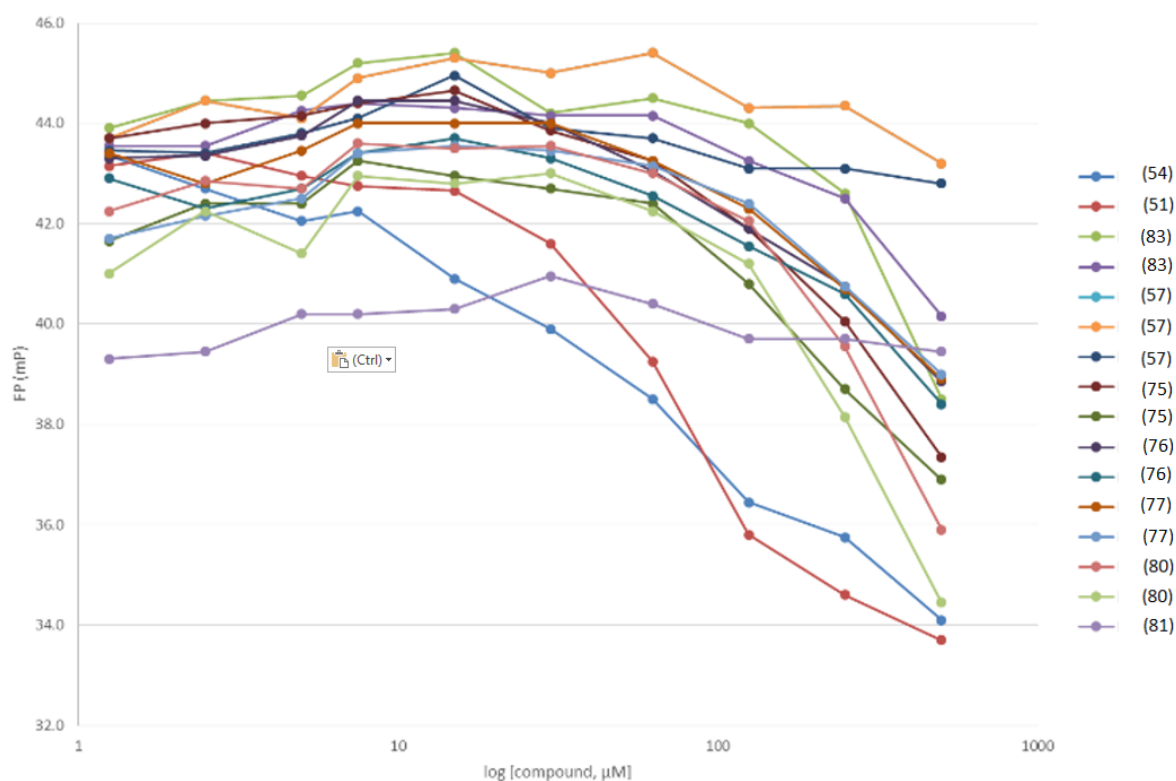


Figure 78; The millipolarisation (mP) versus concentration of compound results for representative analogues of the parent compounds

Most active compounds shown in previous Figure are light orange and dark blue (both compound (57)) and dark purple and green (both compound (83))-Synthesized by Shaun R. Stevens) (Figure 79).

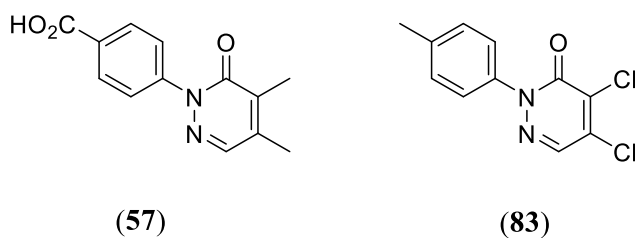


Figure 79; Most active compound after first initial FP assays

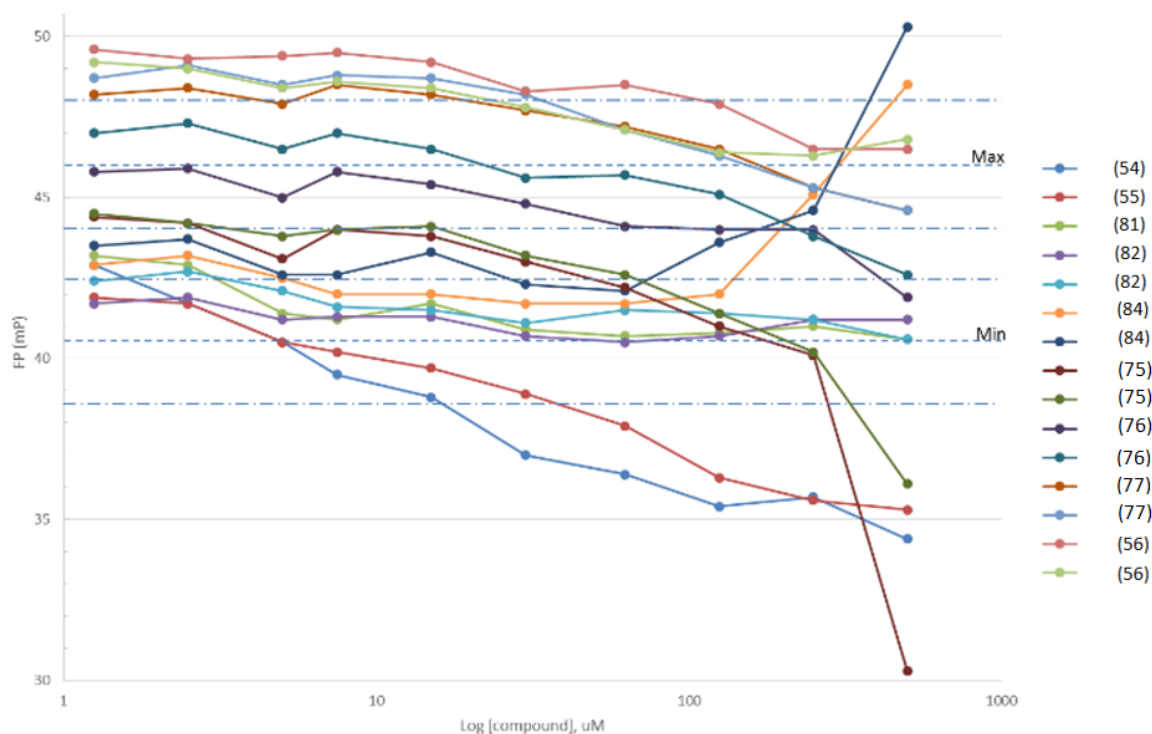


Figure 80; The millipolarisation (mP) versus concentration of compound results for representative analogues of the parent compounds

Most active compounds shown in previous Figure are orange and dark blue (both compound **(84)**-Synthesized by Shaun R. Stevens), and light green and light red (both compound **(56)**) (Figure 81).

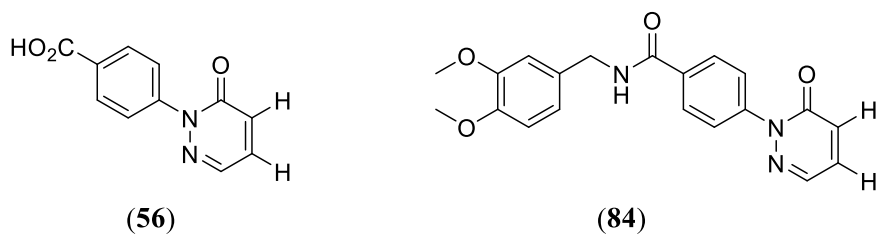


Figure 81; Most active compounds after first initial FP assays

A further optimisation of this assay parameters with the aim of improving the signal-to-noise ratio to a value  $\geq 2.5$  was carried out in order to confirm our hits and to provide us, in combination with other methods ((Isothermal titration calorimetry (ITC), MicroScale thermophoresis (MST), Nuclear Magnetic Resonance Spectroscopy, Differential Scanning Fluorimetry (DSF)) with reliable tools and information to evaluate the activity of the different compounds synthesized and also of the parent compounds.

Unfortunately, none of the methods tried provided us with any useful information and this approach was abandoned. With the hope to obtain additional hit matter a Fragment Based Screening by SPR and NMR was implemented.

## Chapter 11: Fragment-Based Screening by SPR focused on the DNA Binding Domain (DBD)

Additional Hit Finding was done via “Fragment based screening” with the human DNA binding domain (DBD) crystals.

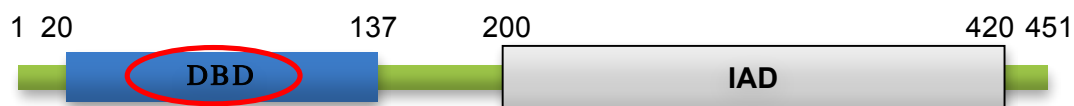


Figure 82, IRF4 domains

The advantage of fragment-based screening (FBS) is the possibility of start an exploration with smaller compounds and build new molecules with more degrees of freedom while not abandoning the “drug-like space”.

The fragmentation of drug leads into smaller fragments, has been used to simplify the computational analysis of ligand binding and to map pharmacophoric elements required for high-affinity binding. The concept of FBS relies on the optimization of each unique interaction in the binding site and subsequent building of a single molecule that should give us a compound with a binding affinity that is the addition of the individual affinities. However, even the most advanced computational algorithms cannot make a perfect prediction of the affinity for each single fragment so, to effectively use fragments in drug design, an experimental method is required to rapidly and reliably screen thousands of low-molecular-mass test compounds for weak (millimolar range) binding to the target protein. Two-dimensional, isotope-edited nuclear magnetic resonance (NMR) spectroscopy is well suited to this purpose and is considered the main method as NMR chemical shifts tend to be sensitive to ligand binding. In our case, a fragment-based screening was undertaken using human DNA binding domain crystals. A subset of the Astex fragment library was used with no hits obtained due to very low ligandability.

After this exploration a new approach, fragment screening by SPR, more efficient towards the detection of very weakly binding hits was undertaken. The SPR experiment is based on the detection of variations in the refractive index upon binding of the ligand to the surface. The variation of the signal is proportional to mass of bound ligand (Figure 83).



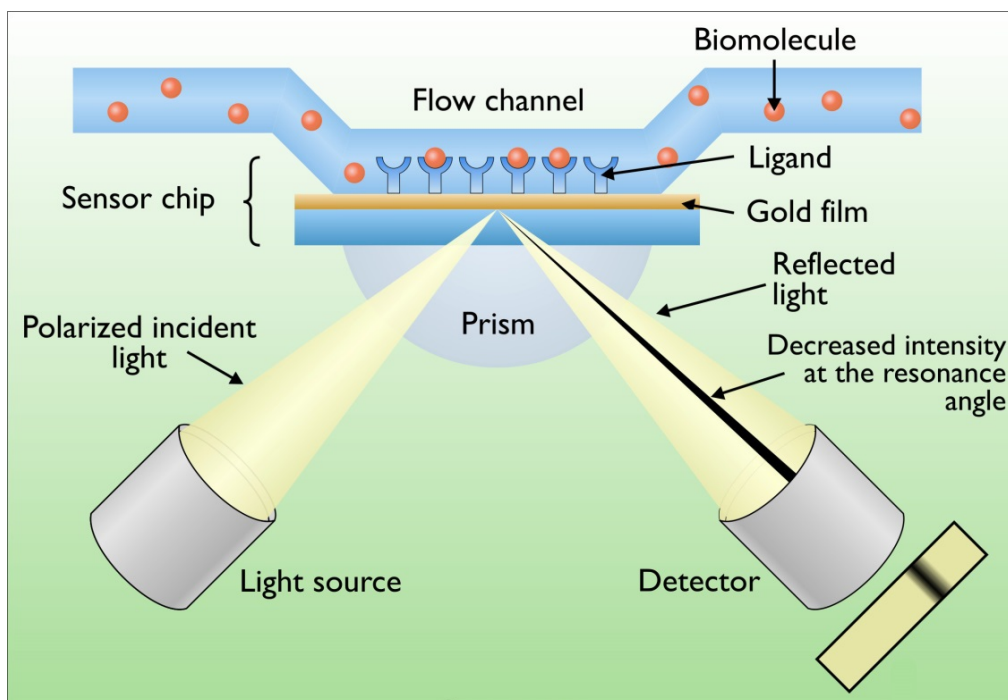


Figure 83; SPR experiment<sup>160</sup>

The SPR experiment was performed at Glasgow University (Figure 80 and Table 3) affording:

- 13 SPR hits out of 1041 fragments (Beatson library) + 3 fragments from Newcastle screened (1.25 %)
- 3 out of 13 SPR hits confirmed by NMR (23 %) (STD and WaterLogsy)

Hits are categorized in:

Hit type	No. SPR hits	NMR validated hits	% validated
Specific (fast $k_{on}/k_{off}$ )	5	2	40
Specific/non-specific (fast $k_{on}/$ slow $k_{off}$ )	2	0	0
Non-specific (slow $k_{on}/k_{off}$ )	6*	1	17

Table 3; SPR and NMR hit finding

The specificity of the ligand towards the target can be measured by the response curves. The “Specific Hit Type” (fast  $k_{on}/k_{off}$ ) is the most promising (Figure 84).

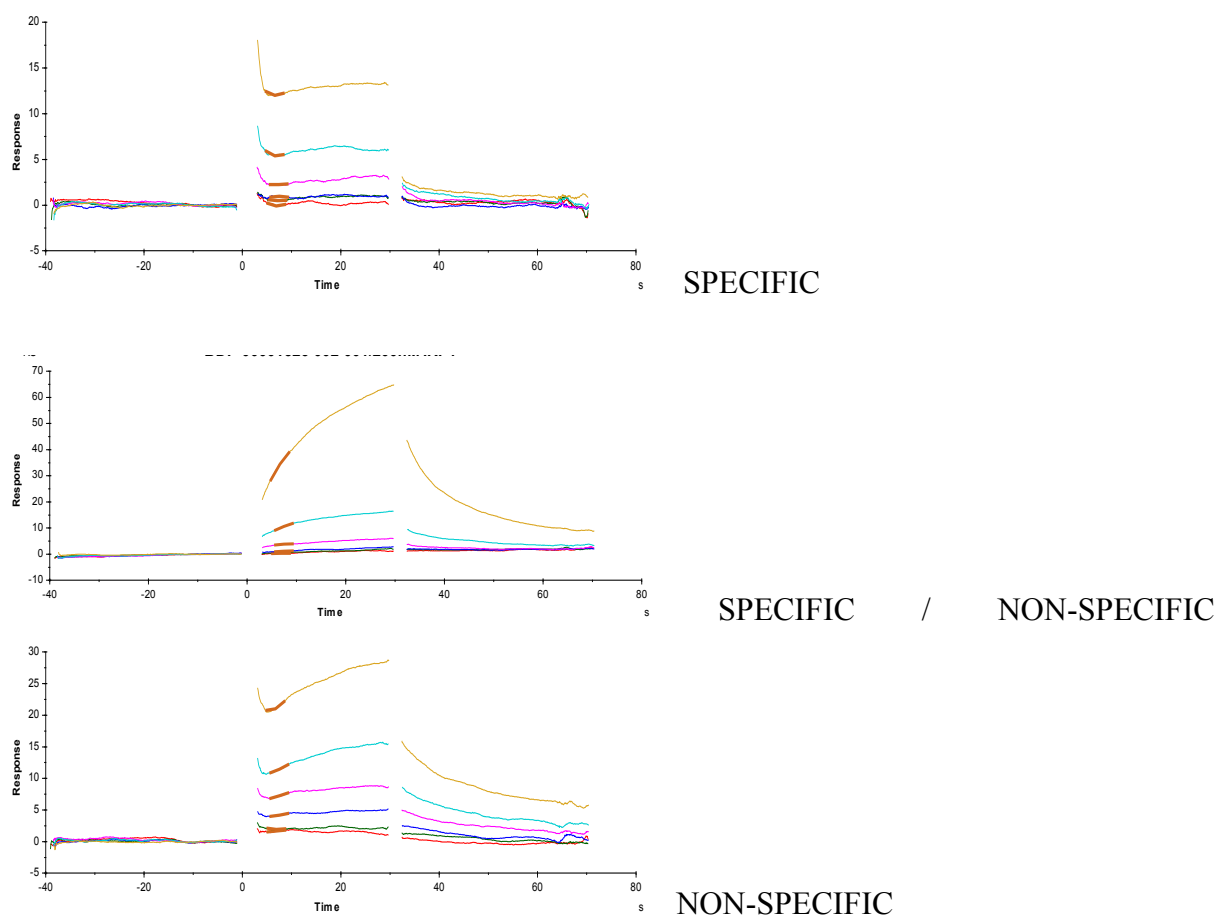


Figure 84; Shapes of general curves for the 3 different hit types by SPR

To confirm the binding site of our compounds and the affinity of those for the DNA binding domain (DBD), the different analogues were soaked into crystals of IRF4 DBD. The different compounds were dispensed directly into crystallisation drops using Echo:

- 25 or 50 nL of 100 mM stock in DMSO
- Final concentration of the compounds  $\geq 5$  to 10 mM
- Soaked for 16 hours – 5 days

The data collected at ESRF or Diamond Light Source is summarized below:

- Resolution range 2 to 2.6 Å
- Auto-processed using Global Phasing software pipeline
- No significant difference density in initial maps

There are several reasons for a lack of electron density; for example:

- Binding site is not present in the crystal
  - Binding pocket occluded by crystal contacts
  - Protein conformation different in crystal than solution
  - Mutations undertaken to form the crystal destroyed the binding site
- Compound does not bind or binds weakly in soak buffer
- Potency/solubility ratio is too low

Despite not having conclusive results a further exploration of the hits obtained was undertaken relying mainly on the SPR analysis.

## Chapter 12: Discussion and Results of the new chemical series

After the fragment-based screening different analogues of the parent compounds identified were synthesized. Herein is reported the synthesis of the parent compounds (**85**) and (**86**) and its analogues (Figure 85).

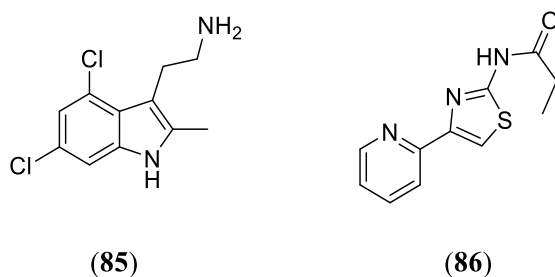


Figure 85; Hits categorized as Tier 1 after fragment-based screen

In both cases the synthesis of the parent compounds and its analogues was undertaken by Stephen H. Hobson, Shaun R. Stevens and me. After obtaining the parent compound (**85**) several analogues were proposed (Figure 86).

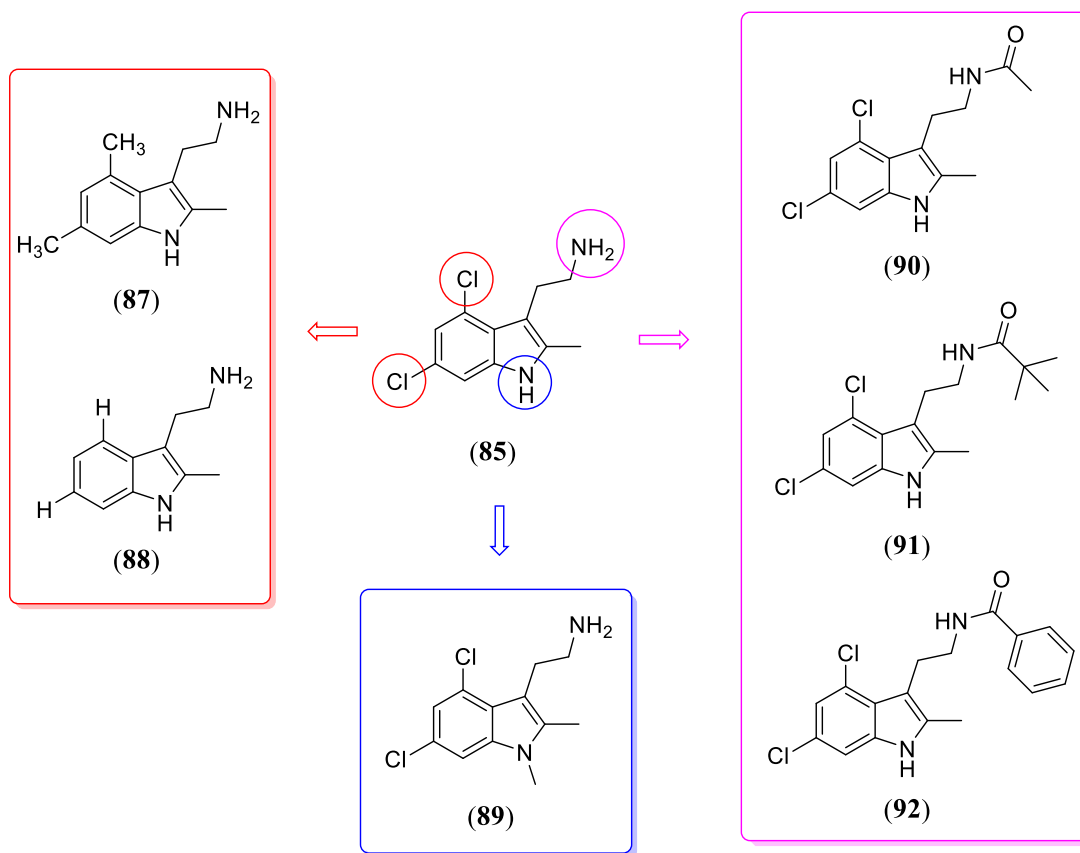
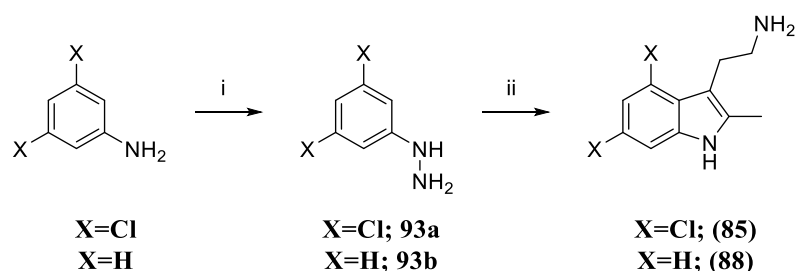


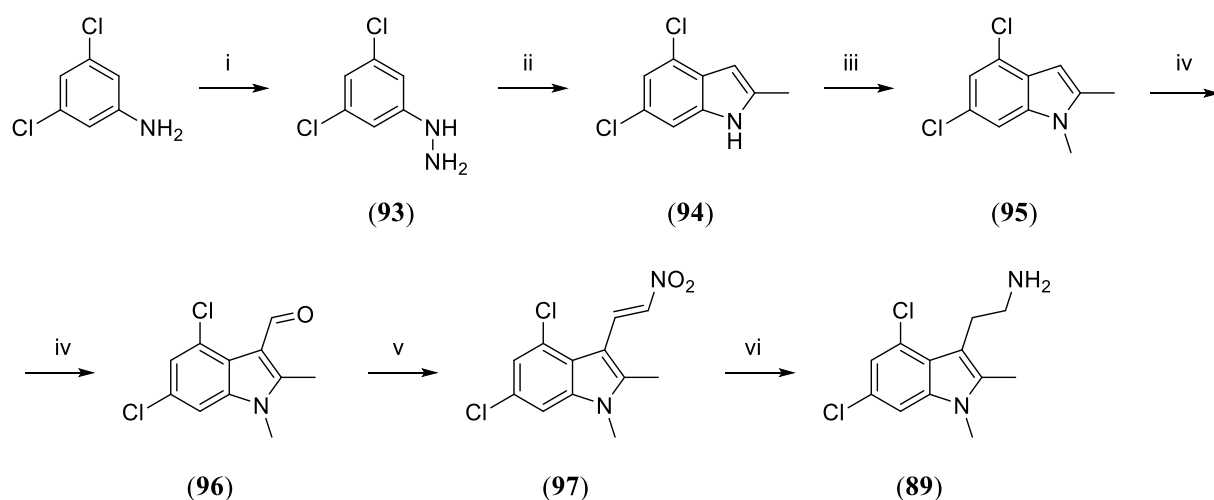
Figure 86; Initial HIT compound and the analogues synthesised

The methodology used to synthesize the parent compound was also used for derivatization (Scheme 29).



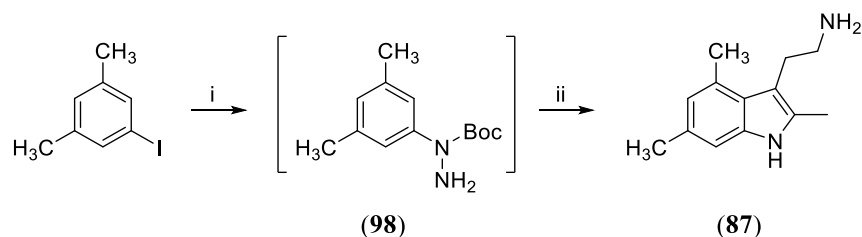
Scheme 29; Methodology used to synthesize the parent compound and its analogues. i) 1)  $\text{NaNO}_2$ , conc.  $\text{HCl}$ ; 2)  $\text{SnCl}_2$ ,  $\text{H}_2\text{O}$ ; ii) 5-Chloro-2-pentanone,  $\text{EtOH}$ ,  $\text{HOAc}$

The aniline is initially dissolved in conc.  $\text{HCl}$  and then sodium nitrite in  $\text{H}_2\text{O}$  is added to the reaction maintaining the temperature of the stirred mixture below  $5^\circ\text{C}$  during the addition. Subsequently the  $\text{SnCl}_2$  in conc.  $\text{HCl}$  is added to form the hydrazine compound **93**. The hydrazine is then treated with 5-chloro-2-pentanone and acetic acid in methanol to give the desired products **(85)**, and **(88)**. Regarding compound **(88)** stability issues were detected and the product was registered but not further tested nor investigated as the molecule was not stable as a solid or in solution. Decomposition was always observed if position 4 of the indole is not substituted with any atom different to hydrogen. In the case of **(89)** a different synthetic route was used due to quaternization of the primary amine after trying to methylate the parent compound **(85)** with iodomethane and using as a base sodium hydride ( $\text{NaH}$  (60%)). The pathway to synthesize **(89)** was via Fischer indole synthesis (Scheme 30).



Scheme 30; Methodology used to synthesize **(89)**; i) 1)  $\text{NaNO}_2$ , (c)  $\text{HCl}$ ; 2)  $\text{SnCl}_2$ ,  $\text{H}_2\text{O}$ ; ii) acetone,  $\text{ZnCl}_2$ ; iii)  $\text{MeI}$ ,  $\text{NaH}$ (60%),  $\text{THF}$ ; iv)  $\text{POCl}_3$ ,  $\text{DMF}$ ; v)  $\text{CH}_3\text{NO}_2$ ,  $\text{NH}_4\text{OAc}$ ; vi)  $\text{LiAlH}_4$ ,  $\text{THF}$

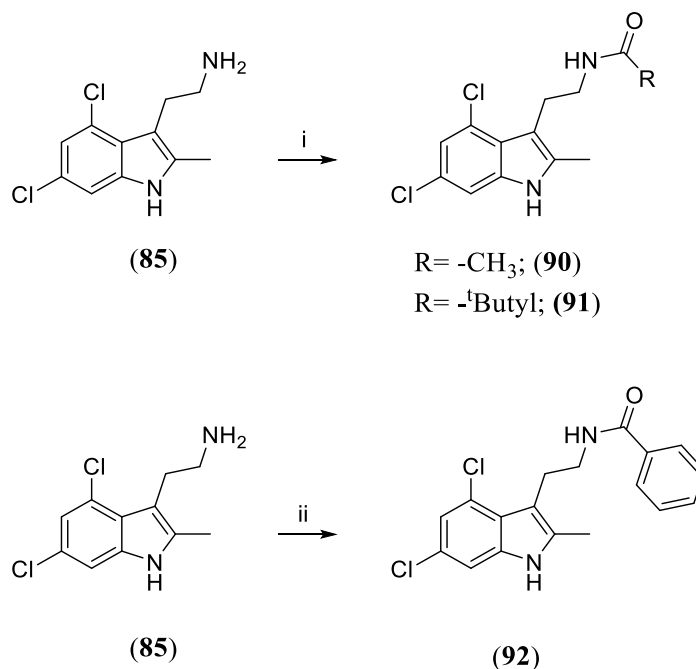
To synthesize compound **(87)** a different synthetic pathway was used due to lack of availability of the commercial hydrazine (Scheme 31).



Scheme 31; Synthesis of **(87)**; i) *t*-Butyl Carbazate, CuI, Cs<sub>2</sub>CO<sub>3</sub>, DMF; ii) AcOH, 5-Chloro-2-pentanone, DMF

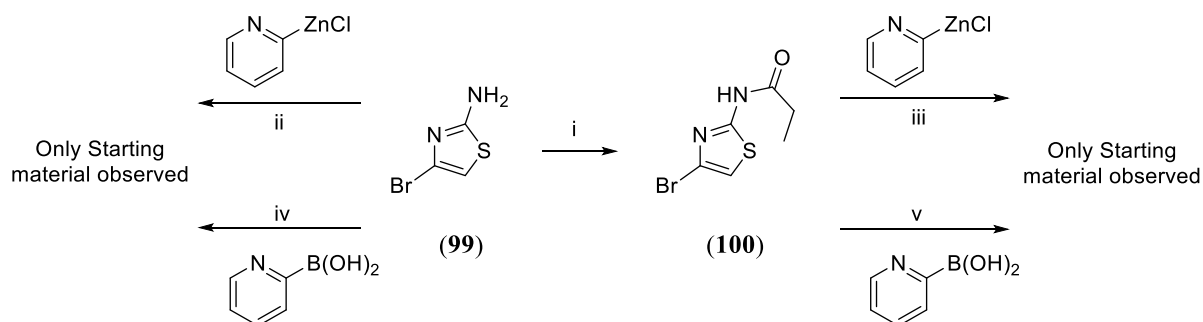
The Boc-hydrazine **(98)** was formed by treatment with *t*-butyl carbazate, copper (I) iodide and Cs<sub>2</sub>CO<sub>3</sub>, the intermediate was directly used without any further purification in the next step to finally synthesize the desired product **(87)**.

For the structure activity relationship exploration of the ethanamine substituent in position 3 of the indole a synthetic pathway similar to the one used for the synthesis of the parent compound **(85)** was used. The parent compound was then used as a common intermediate in a convergent synthesis to obtain the 3 new analogues via amide couplings (Scheme 32).



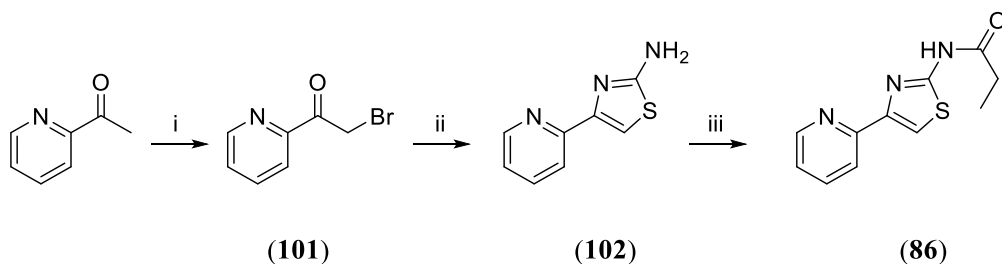
Scheme 32; i) Anhydride, Et<sub>3</sub>N; CH<sub>3</sub>CN/DMF; ii) Benzoyl chloride; Et<sub>3</sub>N; CH<sub>3</sub>CN/DMF

To synthesize compound **(86)** different attempts relying in a 4-bromo-2-aminothiazole intermediate were attempted (Scheme 33).



Scheme 33; i) Propionyl chloride, pyridine, dioxane; ii & iii) nBuLi, ZnCl<sub>2</sub>, THF, Et<sub>2</sub>O; iv & v) Pd(PPh<sub>3</sub>)<sub>4</sub>, Na<sub>2</sub>CO<sub>3</sub>, DMF, H<sub>2</sub>O

Unfortunately, the attempts to synthesize this compound were not successful and a new pathway was followed (Scheme 34). 1-(Pyridin-2-yl)ethan-1-one was treated with bromine in aq. HBr affording compound **(101)**. Condensation of the intermediate with thiourea gave compound **(102)**, which was subsequently treated with propionyl chloride to yield the desired product **(86)**.



Scheme 34; i) Br<sub>2</sub>, aq. HBr; ii) Thiourea, EtOH; iii) Propionyl chloride; pyridine; Dioxane

New analogues were explored by making modifications in the alkyl chain of the amide (right-hand side) and in the pyridinyl moiety (left-hand side) (Figure 87).

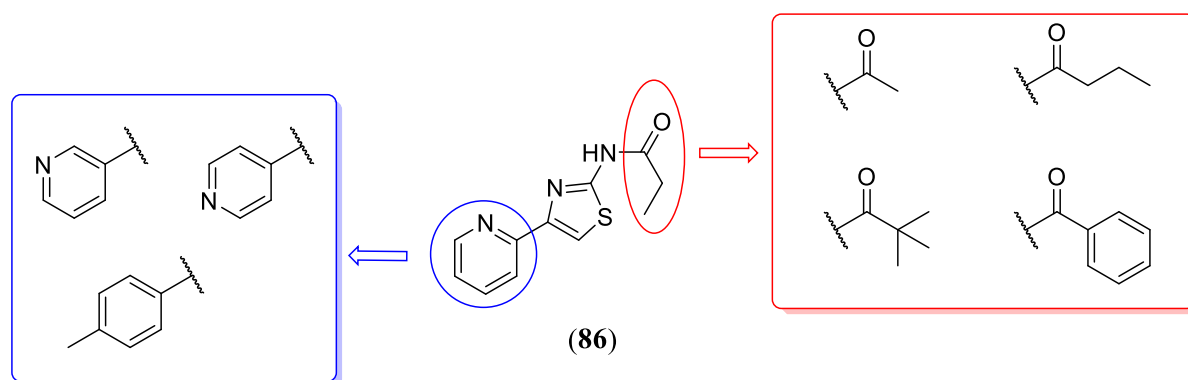


Figure 87; SAR exploration proposed for aminothiazoles

An initial exploration was undertaken by varying the alkyl chain lengths of the amide and by substituting it with a phenyl and a pivaloyl group. The synthesis of the different molecules was done via the same synthetic procedure proposed in Scheme 36. Using compound (**86**) as a general intermediate and different acyl chlorides or anhydrides in the last step of the synthesis gave 4 new analogues in a facile and short synthetic route (Figure 88).

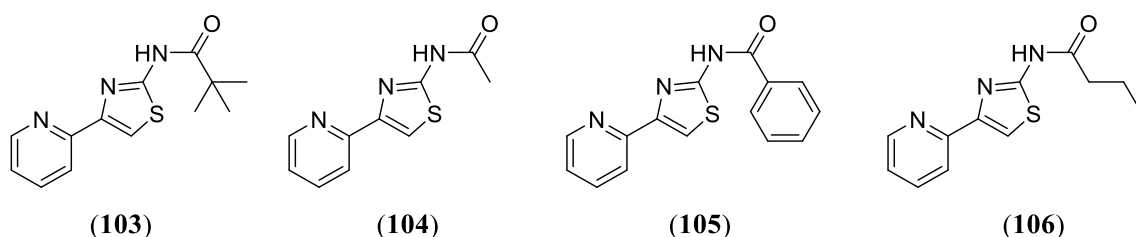
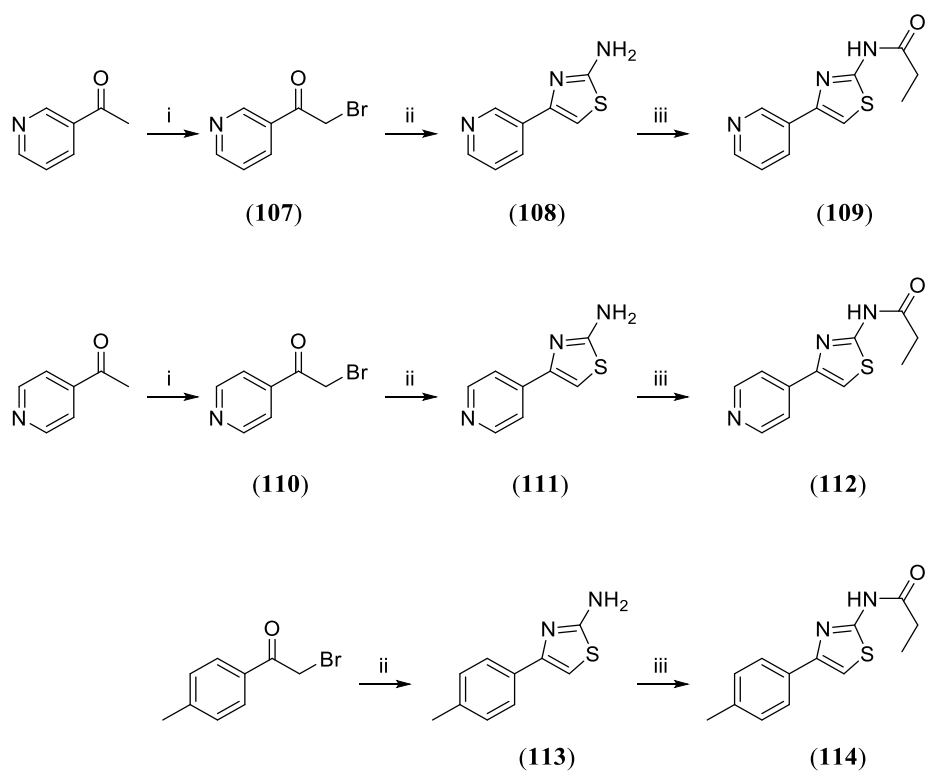


Figure 88; Derivative analogues for aminothiazole series

The exploration of the aminothiazole structure was extended to the pyridinyl substituent, in this case the position of the nitrogen in the ring was modified and the pyridinyl moiety was also substituted for a *para*-tolyl (see Figure 85 left-hand side). The synthesis of the new analogues was undertaken using the same approach shown in Scheme 36. In this case the synthesis was not convergent because no common intermediate can be used for all the targets. A different starting material was used initially in each case to afford the different analogues (Scheme 35).

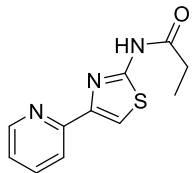
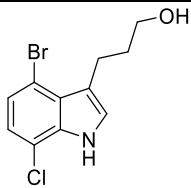
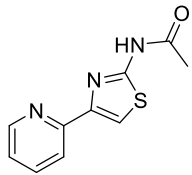
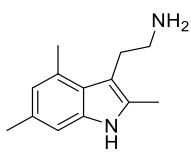
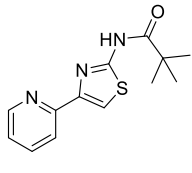
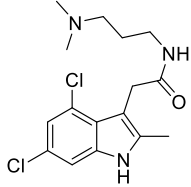
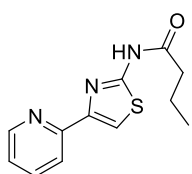
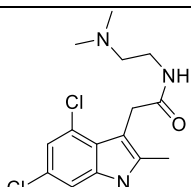
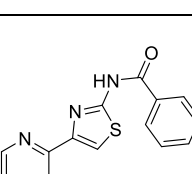
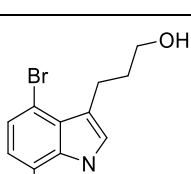
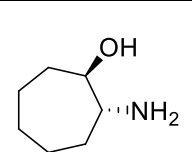
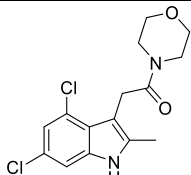


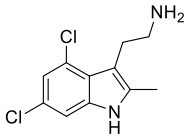
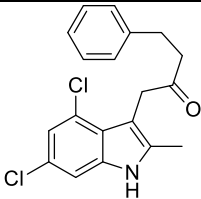
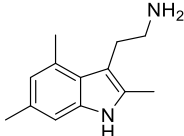
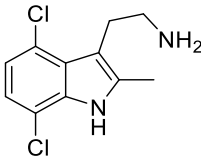
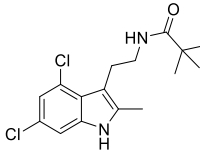
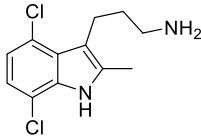
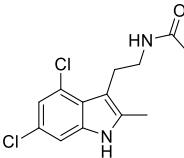
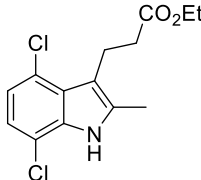
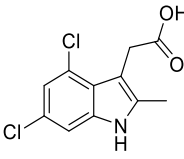
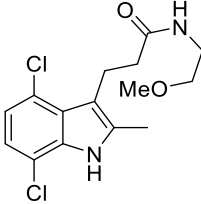
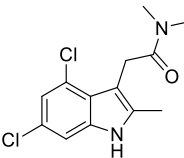
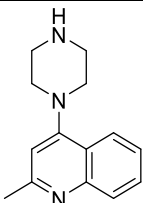
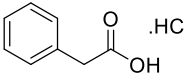
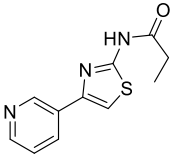
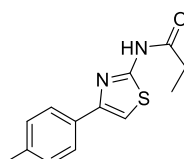


Scheme 35; i)  $\text{Br}_2$ , aq. HBr; ii) Thiourea, EtOH; iii) Propionyl chloride; pyridine; Dioxane

## Chapter 13: Biophysical Assays after FBS and SAR focused on the DNA Binding Domain (DBD)

After the fragment-based screening different analogues of the parent compounds were synthesised (by Shaun R. Stevens (SRS), Stephen Hobson (SH) and me (highlighted with \* in Table 3) or purchased and tested by SPR (Table 4).

ID	Structure	Series	-	ID	Structure	Series	-
<b>86*</b>		Thiazole		SRS-1		4,7-Indole	
<b>104*</b>		Thiazole		<b>87*</b>		4,6-Indole	
<b>103*</b>		Thiazole		SRS-2		4,6-Indole	
<b>106*</b>		Thiazole		SRS-3		4,6-Indole	
<b>105*</b>		Thiazole		SRS-4		4,7-Indole	
-		1-Amino 2-hydroxy cyclo heptane		SRS-5		4,6-Indole	

<b>85*</b>		4,6-Indole		SH-1		4,6-Indole	
<b>87*</b>		4,6-Indole		SH-2		4,7-Indole	
<b>91*</b>		4,6-Indole		SH-3		4,7-Indole	
<b>90*</b>		4,6-Indole		SH-4		4,7-Indole	
SRS-6		4,6-Indole		SH-5		4,7-Indole	
SRS-7		4,6-Indole		-		Quinoline	
-		Pyridine		<b>109*</b>		Thiazole	
<b>114*</b>		Thiazole					

SRS-8		4,6-Indole		SH-8		Amino pyrimidine	
SRS-9		4,6-Indole		SH-9		Indole	
SRS-10		4,7-Indole		112*		Thiazole	
SRS-11		4,7-Indole		SRS-13		4,6-Indole	
SRS-12		4,7-Indole		SRS-14		4,6-Indole	
SH-6		4,6-Indole		SRS-15		4,6-Indole	
SH-7		4,6-Indole		SRS-16		4,7-Indole	

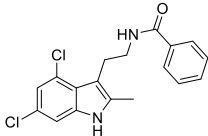
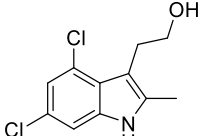
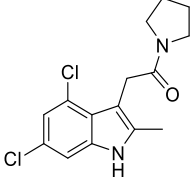
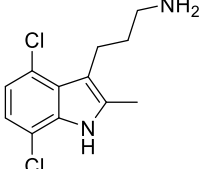
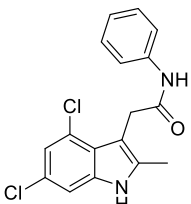
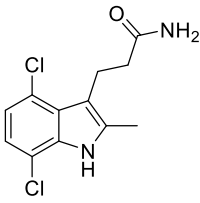
92*		4,6-Indole		SH-10		4,6-Indole	
SRS-17		4,6-Indole		SH-11		4,7-Indole	
SRS-18		4,6-Indole		SRS-19		4,7-Indole	

Table 4; Structure-Activity relationship exploration for the 3 different HITS obtained by SPR  
Fragment based screening (code = **active**, **interference**, **inactive**)

SPR analysis of the different scaffolds obtained by FBS were performed. A summary of the results is shown below:

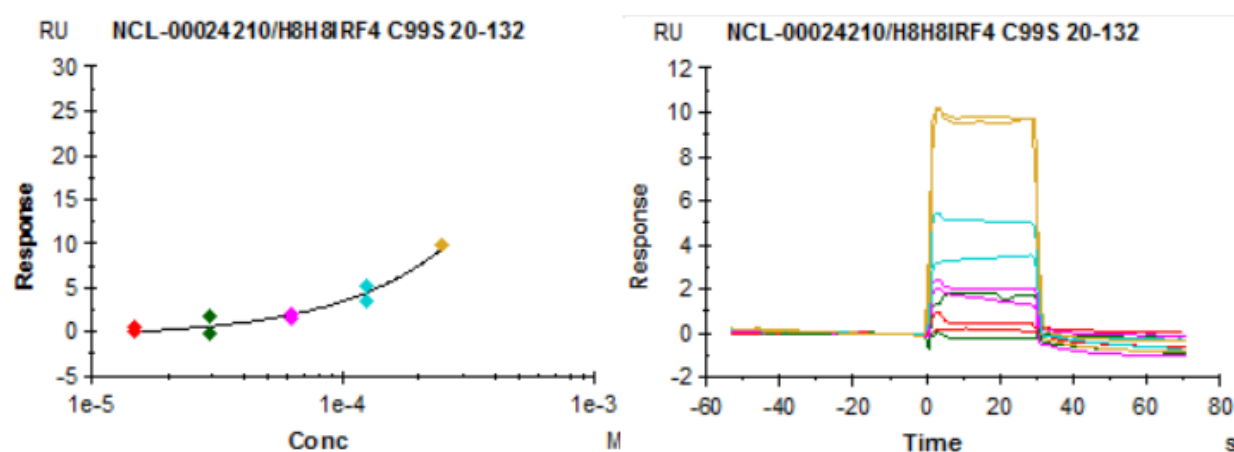


Figure 89; Compound (90) NTA;  $\text{Ni}^{2+}$  2 and 4 His8-His8-IRF4-C99S (BeatsonAnaloguesDR)

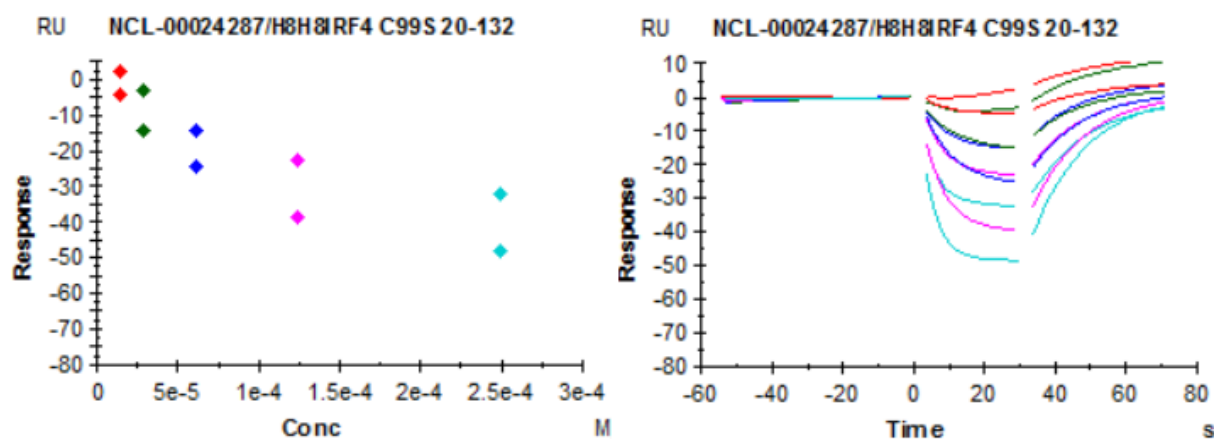


Figure 90; Compound (104) NTA;  $\text{Ni}^{2+}$  2 and 4 His8-His8-IRF4-C99S (BeatsonAnaloguesDR)

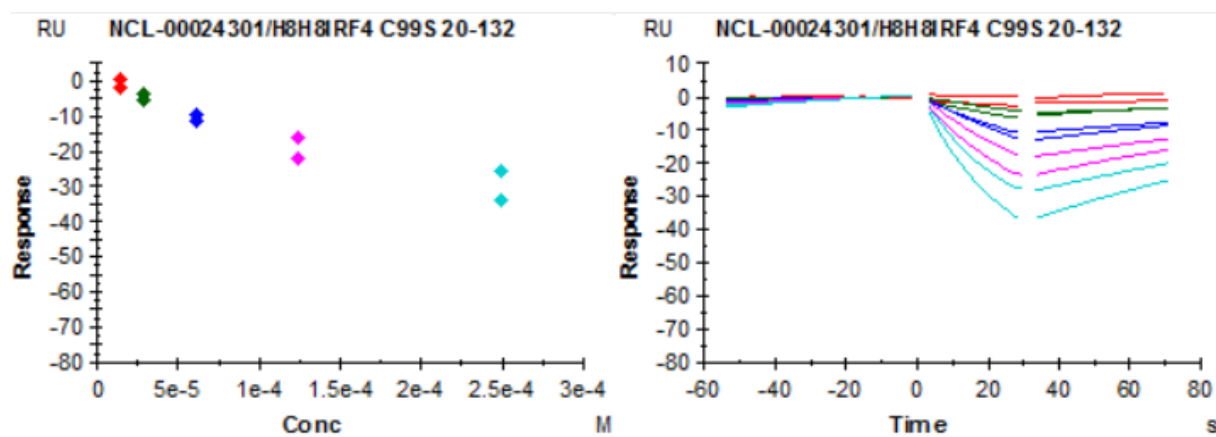


Figure 91; Compound (103) NTA;  $\text{Ni}^{2+}$  2 and 4 His8-His8-IRF4-C99S (BeatsonAnaloguesDR)

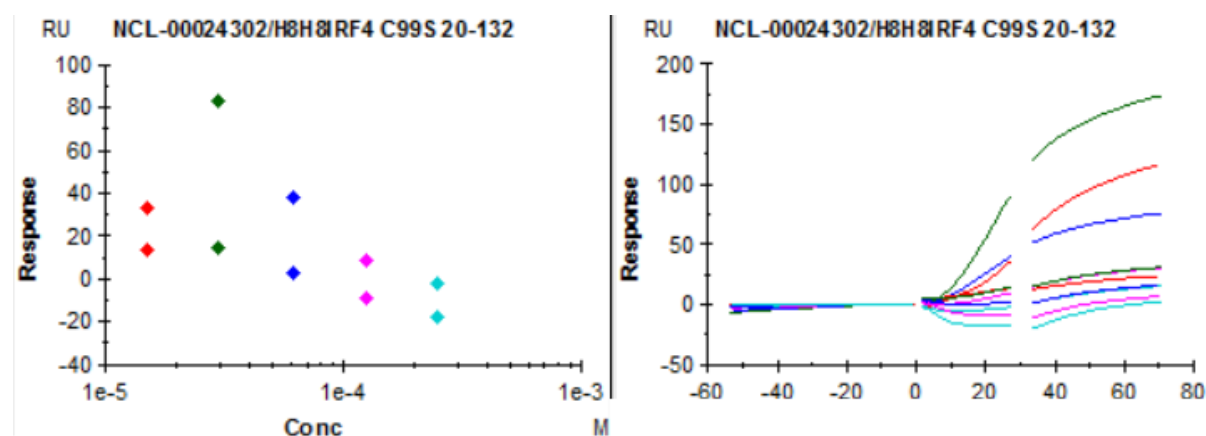


Figure 92; Compound (106) NTA;  $\text{Ni}^{2+}$  2 and 4 His8-His8-IRF4-C99S (BeatsonAnaloguesDR)

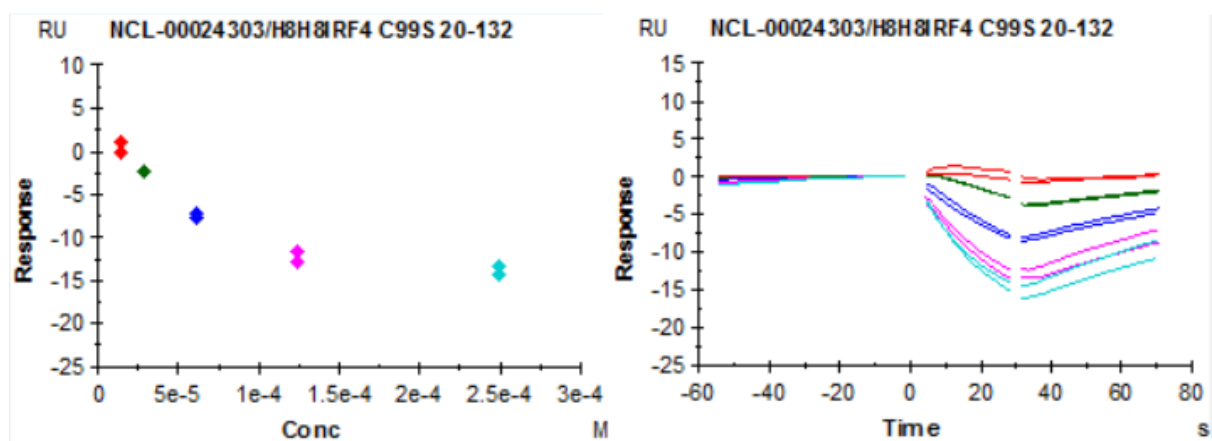


Figure 93; Compound (105) NTA;  $\text{Ni}^{2+}$  2 and 4 His8-His8-IRF4-C99S (BeatsonAnaloguesDR)

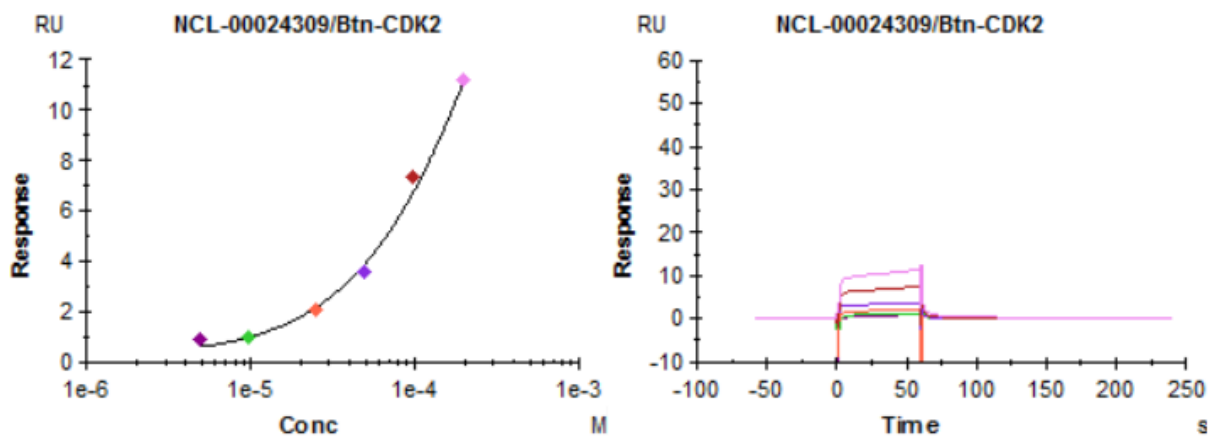


Figure 94; Compound SRS-13 NTA;  $\text{Ni}^{2+}$  2 and 4 His8-His8-IRF4-C99S (BeatsonAnaloguesDR)

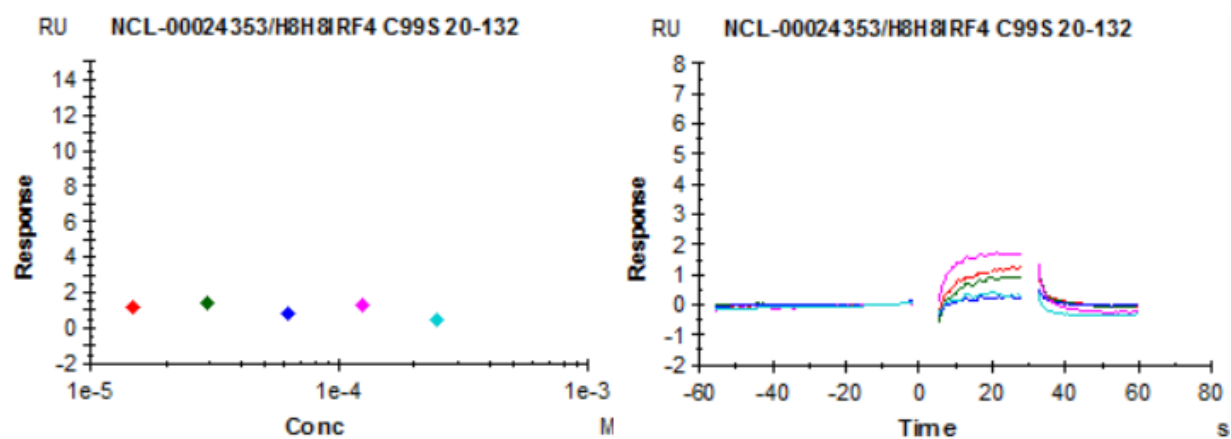


Figure 95; Compound SRS-14 NTA;  $\text{Ni}^{2+}$  2 and 4 His8-His8-IRF4-C99S (BeatsonAnaloguesDR)

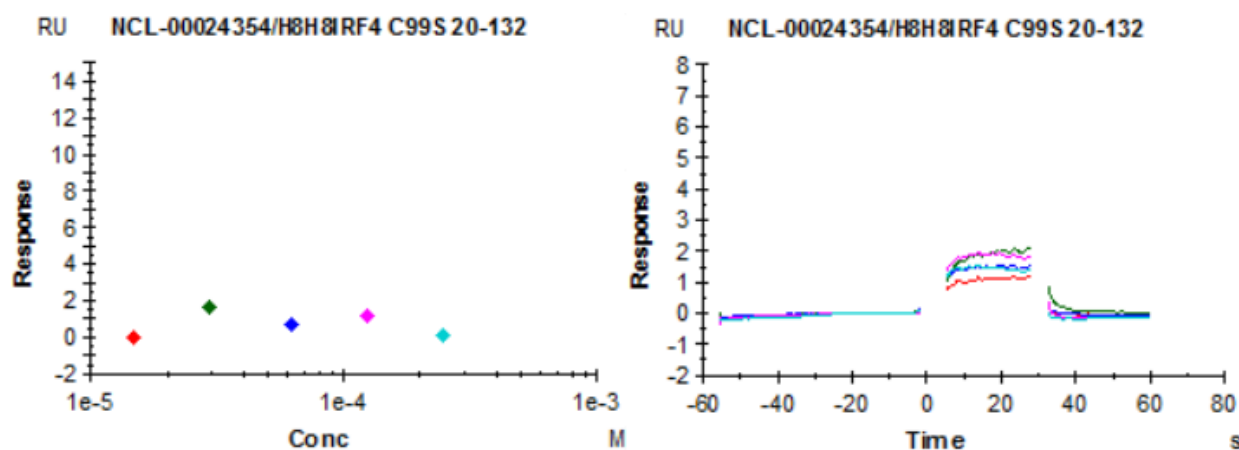


Figure 96; Compound **SRS-15** NTA; Ni<sup>2+</sup> 2 and 4 His8-His8-IRF4-C99S (BeatsonAnaloguesDR)

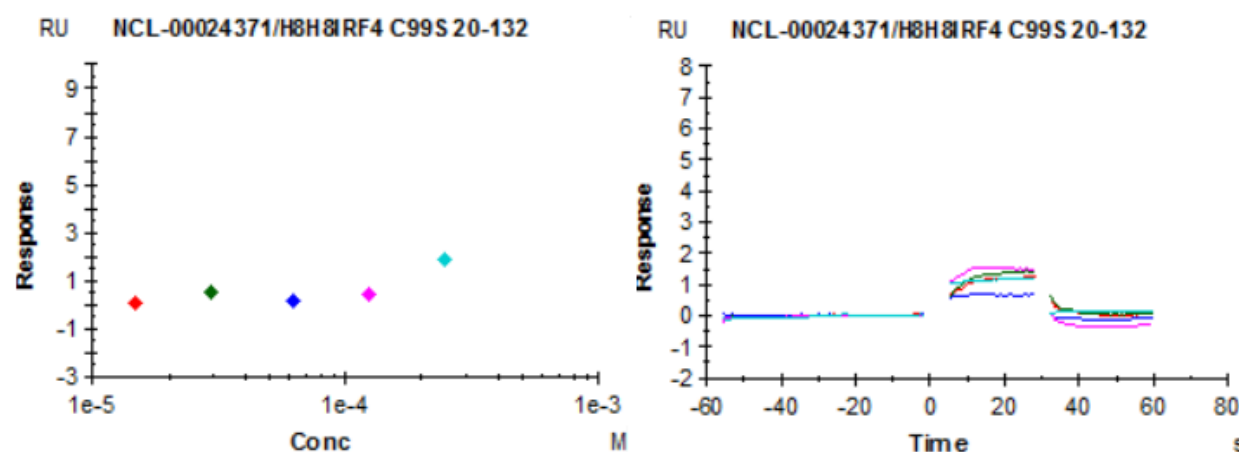


Figure 97; Compound **SRS-16** NTA; Ni<sup>2+</sup> 2 and 4 His8-His8-IRF4-C99S (BeatsonAnaloguesDR)

Unfortunately, the binding of the compounds synthesized was very weak in the best cases (those reported as active in Table 3) as we can see from Figures 89 to 97.



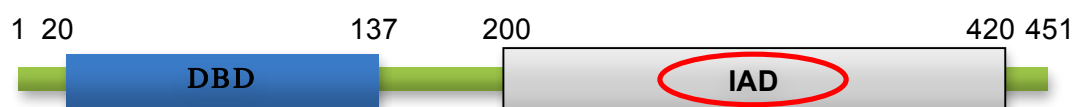
## Chapter 14: Conclusions and Future Work

In conclusion, none of the new compounds synthesized with the most promising SPR results were confirmed by any of the NMR techniques performed (WaterLogsy or STD) or by any biophysical technique. By the terms established as part of our collaboration with Astex Pharma it was concluded that no further work based on these screening results would be financed. From our scientific point of view the lack of confirmation of the hits left us without any solid conclusion.

In the future if a new assay can be developed where the hits obtained can be confirmed it would be interesting to re-test all the compounds synthesized during this thesis. Future efforts must be focused in the development of a functional assay that can allow the scientist an robust development of initial hit matter.

With these results in hand and all the data collected previously, it was decided that no further experiments or structural exploration would be fruitful with any of the hits obtained by HTS or FBS unless a functional assay is developed.

Further experiments with the IRF4 Association Domain (IAD) could be performed in the future. In this approach “Fragment based screening” with human IRF4 association domain (IAD) crystals could be used. The ligandability of this binding site is predicted to be stronger and subsequently easier to confirm by different experiments (SPR, STD, WaterLogsy or X-ray crystal structures), giving us a better opportunity to develop a functional assay. However, the need for a functional assay for the DBD is still there, it would be a great contribution for future research conducted for IRF4 projects. If the assay is developed it should be applied to the previous hit matter disclosed in this thesis for the DBD and also to the analogues synthesized.



At the end of the project our main conclusion was that we could not confirm binding to the DBD of IRF4 by any of the hits found after HTS or FBS or by any of the analogues synthesized and explored in the laboratory during my PhD.

None of the compounds shown here showed significant binding to the DBD of transcription factor IRF4 or in case of binding to the target it could not be confirmed by any technique.

This result was very disappointing from a scientific point of view but new questions came to our minds:

- “Could we save time and money to ourselves in future projects by NOT spending more time than necessary in undruggable targets?”
- “Could we map the interactions of a ligand to the protein in a facile and fast way?”

It is often not apparent from the outset whether it will be possible for a small molecule (or subset of molecules) to bind to a given site with sufficient affinity to modulate its function. As a result, many drug discovery projects are initiated against chemically intractable targets.

To try to assess druggability in a facile, faster and reliable way a new method is described in the following section of this thesis: **FragLites**

## **Chapter 15: FragLites. Developing new approaches to assess druggability and hit finding at earlier stages**

Drug discovery is characterized, among other things, by drug attrition, which could be described as the high failure rates associated with pharmaceutical development. Despite of the technological advances and increase in research and development efforts, the number of truly innovative medicines approved (1<sup>st</sup> in class) is not increasing accordingly. Low productivity of NMEs can be tackled in many ways, from an organizational point of view or a scientific point of view.

From an organizational point of view, it was theorized that an increased workload in the pre-clinical stages would lead to more drugs being successful in the clinical stages. The way companies have improved productivity is related to an increase in the chances of success in the clinic by focussing on better rationale for targets and compounds.

### **15.1 Improvements in the scientific methodology to speed-up the drug discovery process**

In the last decades (especially since the 1990s) the genomics have dominated the way new biological targets (usually proteins) are discovered to be the key regulators in the development of a disease. The modulation or suppression of the activity of these targets gives us a rationale base for drug discovery. This approach takes us to a target-centric strategy, where different biological hypotheses can be tested and used as starting points for molecule identification.<sup>161</sup>

Several methods have been developed to find drugs that interact with the desired biological target. One of the most widely used is High-throughput screening (HTS). This method has been tremendously potentiated by robotics (Figure 98) to conduct millions of chemical tests where a researcher can quickly identify active compounds that interact with a particular protein (called hits).

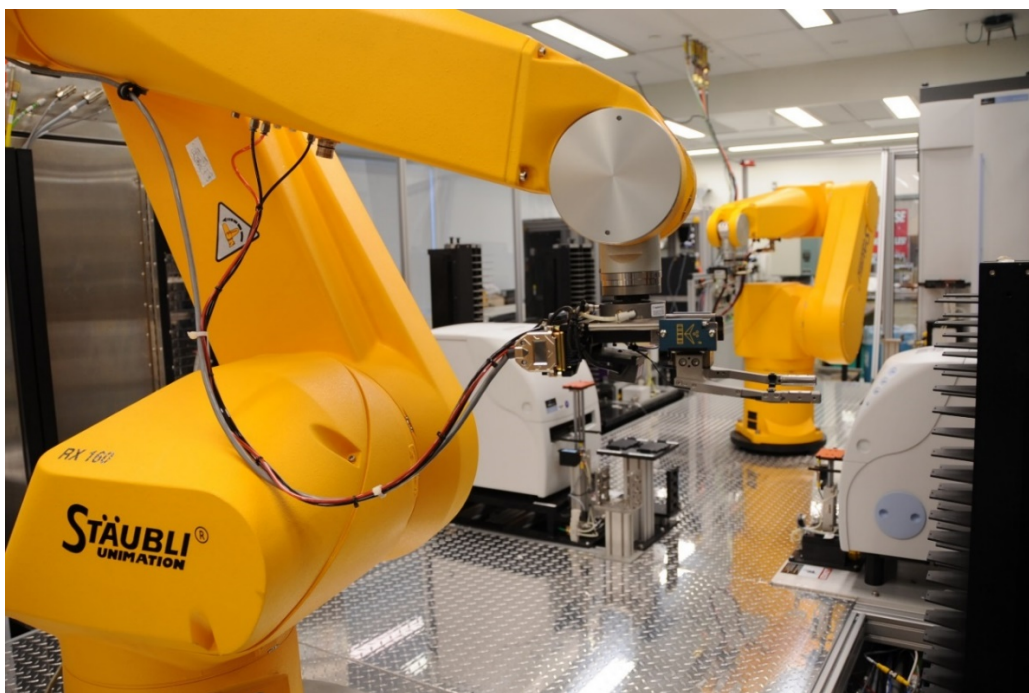


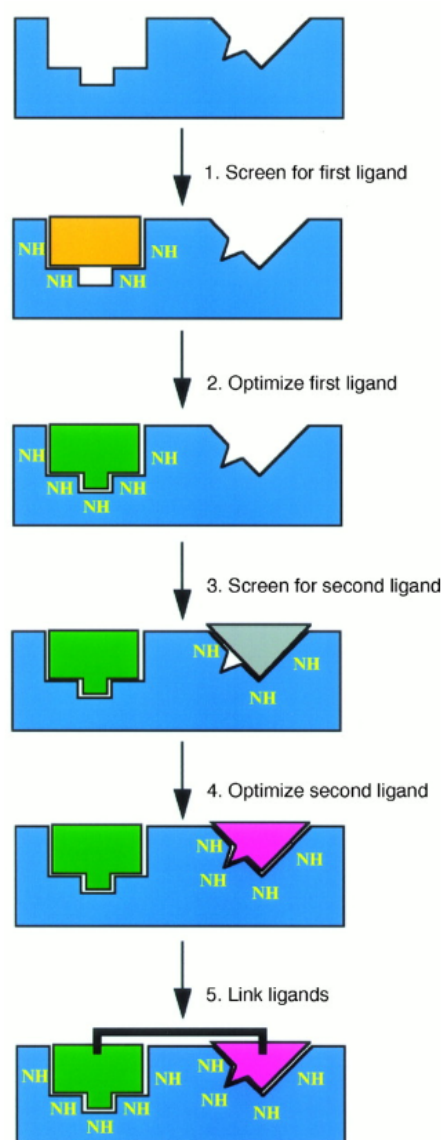
Figure 98; A robot arm handles multiple assay plates

Recent advances in automation and drop-based microfluidics allowed this technique to effectively test 100 million reactions in 10 hours but these methods are still very expensive to implement. However, traditional HTS is increasingly utilized in academia thanks to the growing availability of compound sets. At the same time, many companies specialized in these services offer the possibility of outsourcing HTS campaigns, making HTS the leading paradigm of small molecule drug discovery.

During the 1990s several flaws associated with HTS campaigns were known, the most important liability was the tendency to obtain as hits with very high lipophilicity and drug-unlike properties. To avoid these issues drug-like criteria (like the Lipinski rule of five<sup>162</sup>) was developed for the design of the screening libraries. Nowadays, target-focused compound collections can be found if we already have a starting point. Screening libraries can also be *in silico* designed.

However, the lack of success of HTS for some specific targets, gave a steady impulse to the creation of new techniques like Fragment Based Drug Discovery (FBDD), which development started with the “SAR by NMR” by the Abbott group in 1996.<sup>163</sup> In this paper, it is described how a linked-fragment approach, where ligands are treated as building blocks with independent binding to different parts of the target, can yield a final drug with optimized

affinity for the protein. Binding is determined by the observation of  $^{15}\text{N}$ - $^1\text{H}$ -amide chemical shift changes in a 2D  $^{15}\text{N}$ -heteronuclear single-quantum correlation ( $^{15}\text{N}$ -HSQC) spectra. Once a lead molecule is identified the binding is optimized to its site, then the process is repeated for a ligand in an adjacent pocket. Then the two fragments are linked together with the right orientation so the affinity of the new (and larger) ligand should be improved (Scheme 36).



Scheme 36; Outline of the SAR by NMR method (Reprinted with permission from AAAS)

Over the last 20 years we have seen increasing acceptance and use of the fragment-based approach. The key of FBDD is that it does not only rely on Ligand Efficiency as the fundamental key for binding but also in the efficient fit to the binding site of a target by the ligand.

In the early 2000's emerged the rule of three.<sup>164</sup> A useful rule to define a fragment for the construction of libraries for lead generation, fragment hits must have:

1. Molecular weight of  $< 300$
2. ClogP  $< 3$
3. Hydrogen bond donors and acceptors  $< 3$

To study the development of fragment-based drug discovery and the impact on drug discovery Johnson *et al.* have published a mini-perspective article summarizing hit to lead development (2015-2017), nature of the screening methods used, and the target classes involved in these hits. An analysis of the distribution of target classes thanks to these articles, shows that there was an increase in the development of antagonists of protein-protein interactions (specially for bromodomains) (Figure 99), which represent a target with X-ray crystallography availability. About three quarters of the campaigns target enzymes, and around one quarter of the examples specifically targeted kinases. In the last three years, only one GPCR and no ion channels have been reported. These data reflect the difficulty in developing FBDD without high resolution X-ray structures.

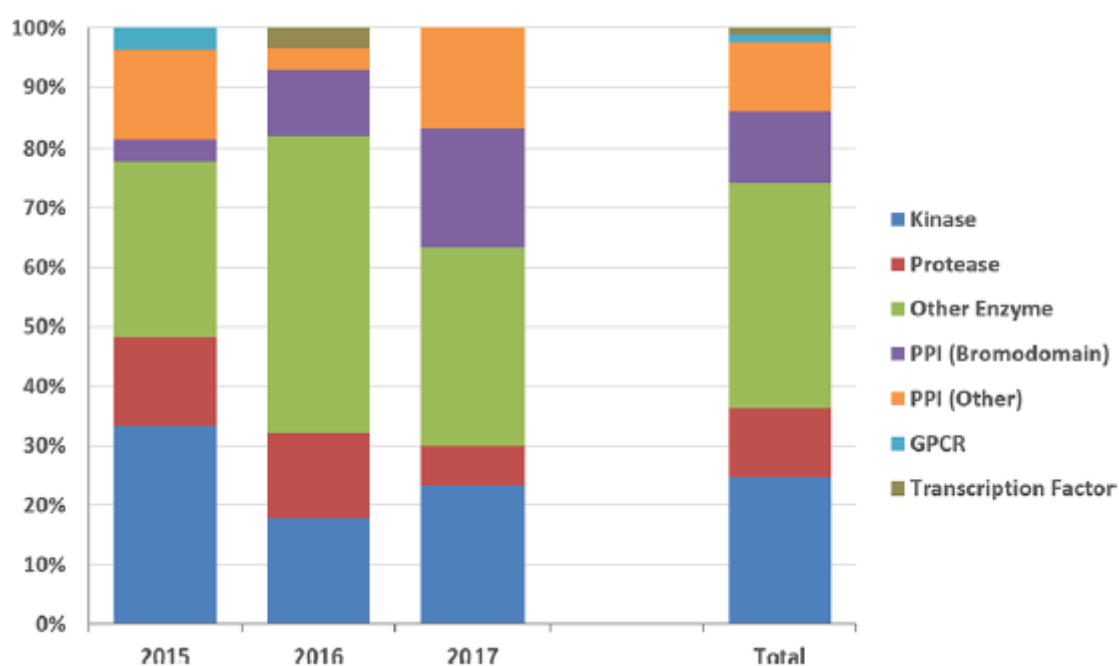


Figure 99; Distribution of target class for the F2L examples

The choice of screening method is highly relevant. Biochemical assays are the most commonly used screening technique (5 out of 7 kinases, 8 out of 11 for the rest of enzymes, only 2 out of 11 protein-protein interactions) (Figure 100). In the later years there was an increase in the number of virtual screens, probably derived from the improvements in the techniques and the increased availability of structural information. HTS can also be integrated with FBDD (deconstruction of HTS hits), and these approaches are potentially complementary. We can use an initial fragment collection for screening (FBDD), then analyse the results and rationalize the selection of compounds for a new screening by HTS containing the appropriate sub-structural features.<sup>165</sup> In addition to the rational deconstruction of hits derived from HTS, a fragment obtained by FBDD can be merged with a hit obtained by HTS.

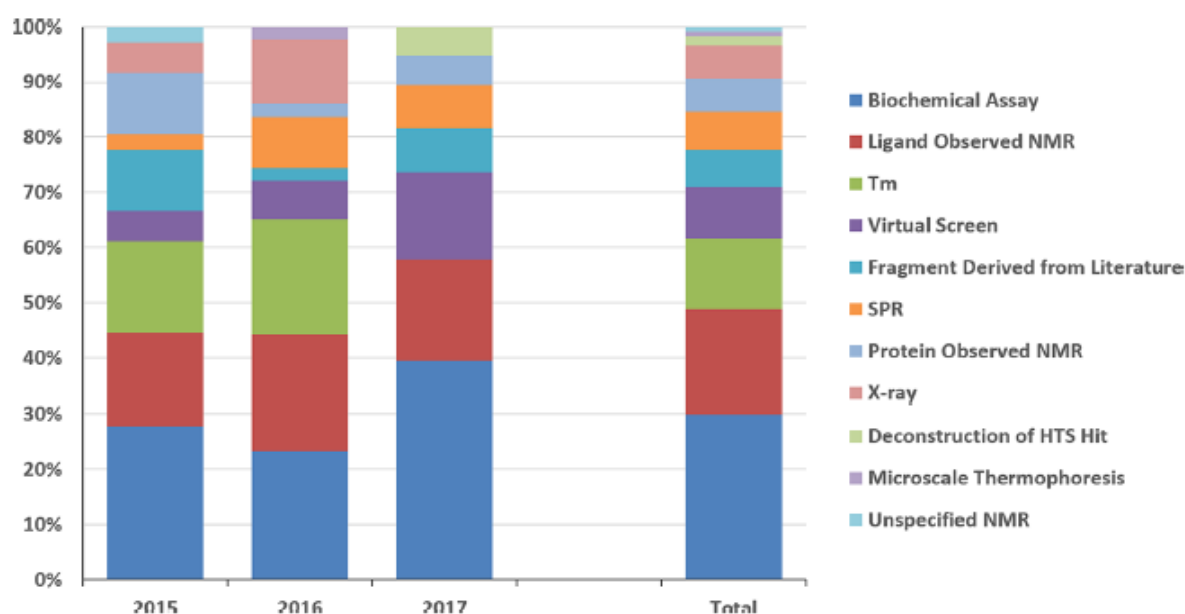


Figure 100; Distribution of screening method for the F2L examples along with the total

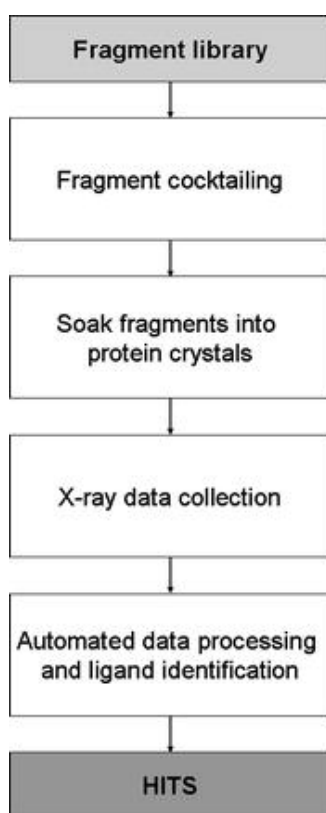
The articles published by Johnson *et al.* state the importance of structural information for fragment optimization. In 21 out of 30 campaigns a crystal structure was available.

X-Ray crystallography is the most sensitive of the biophysical techniques used for fragment-screening experiments. The lack of theoretical lower limit on the affinity of fragments detectable is the main advantage, with the practical limitations being compound solubility and crystal robustness. This provides us reliable detection of compounds with a dissociation constant  $K_d \leq 5$  mM. Furthermore, the use of crystallography as a screening technique has

other advantages like specific structural information on the interaction between fragment hit and target at the earliest stage, providing fast and efficient assessment of hits in terms of their medicinal chemistry tractability and in terms of synthetic vectors that can help in the optimization.

A disadvantage is the low throughput of X-ray crystallography. Different methods have been described allowing the power of X-ray based screening to be fulfilled as a highly viable component of drug-discovery. One of the more recent methods was developed by Astex and is called “Pyramid™”. In this approach, high throughput crystallography is coupled with other biophysical techniques, such as NMR, resulting in an integrated fragment-based discovery platform.<sup>166</sup>

The Pyramid screening idea is based on “automated protocols for rapid X-ray data collection”. The method requires the soaking of crystals with fragments of interest, followed by X-ray data collection, processing, and refinement of the ligand-free complex to potentially reveal the difference in electron density associated with the bound ligand. The electron density is then interpreted, fitted, and the complex further refined to give the final protein–ligand structure (Scheme 37).



Scheme 37; Flow-chart for a crystallographic fragment-screening experiment



The Astex's Core Fragment Set was created as a general-purpose library (1,000 fragments approximately), which aims to cover wide chemical space and be suitable for screening against different targets. The rationale for the library creation is based on the idea that "drug-fragment space" can be effectively sampled with a relatively small number of compounds based on scaffolds and functional groups commonly found in drug molecules.

Fragments are screened by soaking the ligand of interest into pre-formed protein crystals, with a high concentration solution of ligand in mother liquor for enough time. Another possibility is to co-crystallize the complex *in situ*. Effectively, a separate crystallization method is needed for each compound and the process gets even more complicated if for each compound, different crystallization conditions are needed.

Unfortunately, this method will give us a large library of compounds to be tested and experimental methods such as fragment screening by X-ray crystallography cocktail are still time consuming. Moreover, to adequately cover chemical space, it is required the testing of a sufficient numbers of compounds, such that a druggability assessment using conventional fragments requires a comparable investment to running a full fragment screen. For this reason, there is scope to further improve FBDD screening, especially regarding the fast acquisition of reliable crystal structures.

## Chapter 16: Project Rationale

Since the anomalous signals from native atoms are relatively weak, it is normally advantageous to add heavier atoms to protein crystals either through incorporation during protein expression,<sup>167</sup> co-crystallization as a binding partner<sup>168</sup> or soaking into previously formed crystals.<sup>169</sup> During an X-ray crystallographic small-molecule fragment screening campaign targeting HIV-1 reverse transcriptase bound to a non-nucleoside reverse transcriptase (NNRTI) drug, (rilpivirine), 4-bromopyrazole was found to bind to 15 sites and 4-iodopyrazole in 21 (Figure 101).

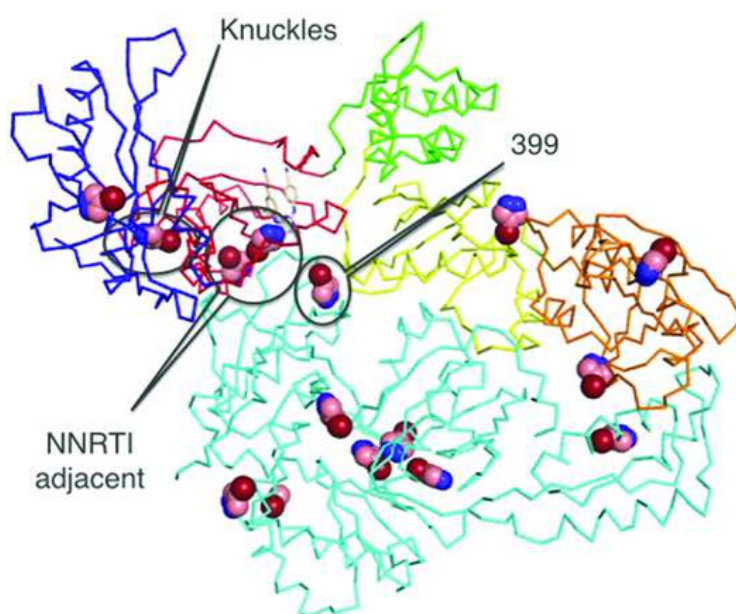
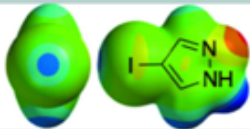
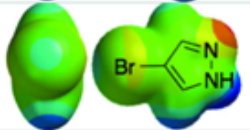
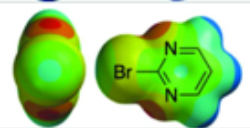
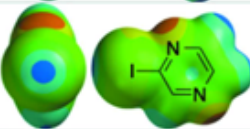
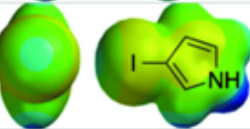
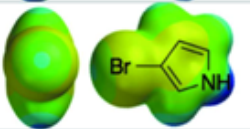


Figure 101; Locations of 4-bromopyrazole are shown as light pink spheres

In general, to have sufficient affinity, size and shape to be detected by X-ray crystallography, compounds typically need to have molecular weights  $> 200$ . It has been theorized that the promiscuous nature and detection of molecules like 4-bromopyrazole or 4-iodopyrazole is due to electrostatics. Density functional theory calculations revealed that the electronegative nitrogen atoms in the pyrazole ring enhance the polarization of the halogen atom in both, negative and positive charge of the 'sigma hole' opposite to the C—X bond (Table 5).

Compound name	Electrostatics	No. of sites on RT
4-Iodopyrazole		21
4-Bromopyrazole		15
2-Bromopyrimidine		4
Iodopyrazine		2
3-Iodopyrrole		0
3-Bromopyrrole		0



  
 -200                      kJ mol<sup>-1</sup>                      +200

Table 5; Computed electrostatic surfaces for 4-Iodopyrazole, 4-Bromopyrazole and related analogues

The promiscuous binding of 4-bromopyrazole and 4-iodopyrazole allows the ordered binding of Br and I atoms, respectively, at many sites around a target protein and both bromine and iodine scatter X-rays with strong anomalous signal when appropriate wavelengths are used. While in typical non-crystallographic screenings the present promiscuous binding would be undetectable, in the case of crystallographic fragment screening it is possible to see how the compound binds to the target and in how many sites across the protein (Figure 102).

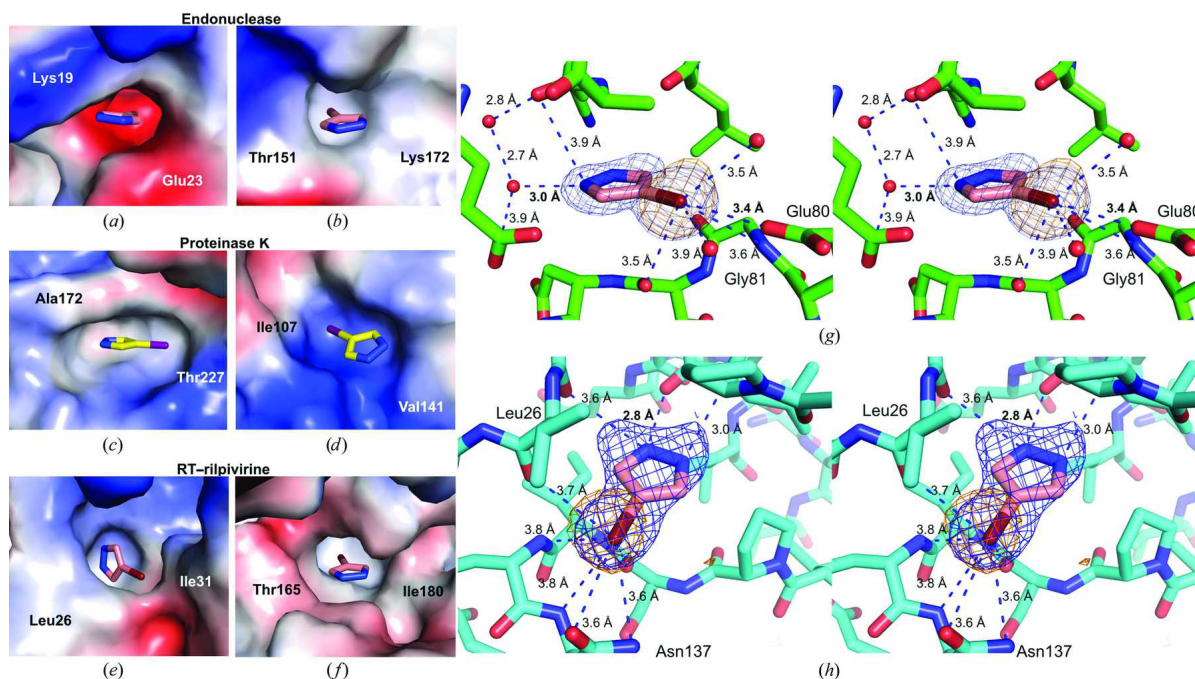


Figure 102; 4-Bromopyrazole and 4-iodopyrazole binding sites. (a)-(f) Two binding sites with the clearest electron density for the compounds are depicted for each protein: (a, b) influenza endonuclease, (c, d) proteinase K and (e, f) RT-rilpivirine. Electrostatic potential surfaces are shown. Potential electrostatic interactions are depicted as blue dashes with interatomic distances shown

The unambiguous detection of this anomaly makes possible the detection of the exact location of these small fragments. Usually, a fragment is considered a molecule with a molecular weight of  $< 300$ , and when the libraries are designed in order to be able to detect our fragments the molecular weight usually varies between 200-300. 4-Bromopyrazole (molecular weight is 144) is smaller than this. The possibility of detecting these small fragments in an electron density map gives us the possibility to explore more chemical space with smaller libraries, and directly test small binding motifs in our targets.

A structure-based assessment of druggability could be achieved more efficiently and provide guidance into the likely pharmacophores and chemotypes that would bind to the target protein. In addition, the ability to map a protein for all potential sites of ligand interaction could be provided, identifying opportunities for allosteric inhibitors, with great potential to tackle resistance mechanisms, both alone and in combination with orthosteric ligands. It would be advantageous to develop a method that can quickly assess the druggability of a target in an efficient way, without consuming a large amount of time and resources.

To generate evidence that these problems can be addressed, a set of very simple fragments expressing common pharmacophoric interacting groups was designed to identify productive binding sites using X-ray crystallography. The set consists of 31 fragments, which express the combination of two binding hydrogen bonding groups (donors and acceptors) separated by 1 to 4 bonds and incorporating a bromine atom for visibility in the electron density map. Also, the combination of two polar groups should confer water solubility and allow high concentrations of these molecules for crystallography studies (Table 6).

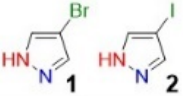

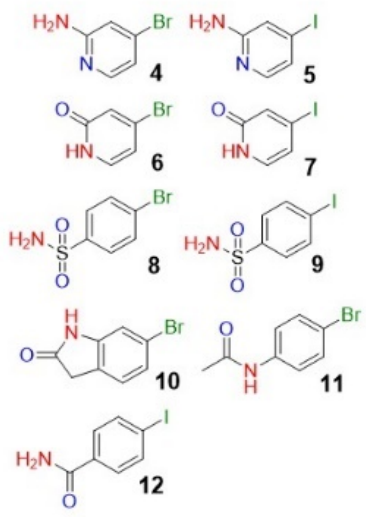
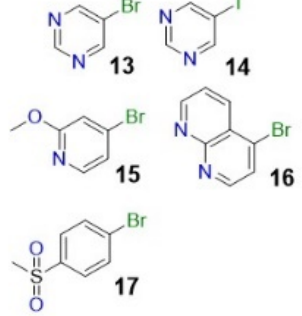

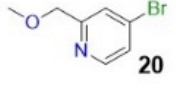

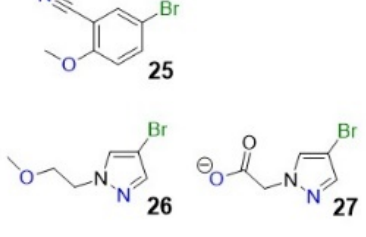
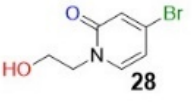

Connectivity	Donor / acceptor	Acceptor / acceptor
1	 1 2	 3
2	 4 5 6 7 8 9 10 11 12	 13 14 15 16 17
3	 18 19	 20
4	 21 22 23 24	 25 26 27
5	 28	 29 30 31

Table 6; Design of **fragments** and matching with pharmacophore doublet

## Chapter 17: Results and Discussion

To test the library and to validate our concept, initial testing was done in a protein for which druggability is established and potent ligands are known. Cyclin-dependant kinase 2 (CDK2) is a protein that has been explored recently in our group by Celine Cano *et al.*<sup>170</sup> and the existence of several allosteric sites on top of the orthosteric has been previously shown.<sup>171</sup>

Each of the FragLites from our library was soaked into CDK2 protein. The anomalous signal generated from our fragments allowed us to detect 9 of the 31 fragments in 6 different sites. Fragments **1**, **2**, **7**, **13**, **14**, **16** and bound the orthosteric site with fragment **31** binding in an adjacent pocket. The orthosteric site was successfully identified thanks to the highest frequency of binding. The Fraglites library also identified known and new allosteric sites. Of the 5 detected allosteric sites, fragment **4** interacts in the lower region of the C-terminal lobe on the C-helix side. Another allosteric site was observed with fragment **6** at the very bottom of the C-lobe. The other sites were all identified by fragment **31** at the N-terminal lobe below the hinge region and in the pocket close to the cyclin binding region (Figure 103).

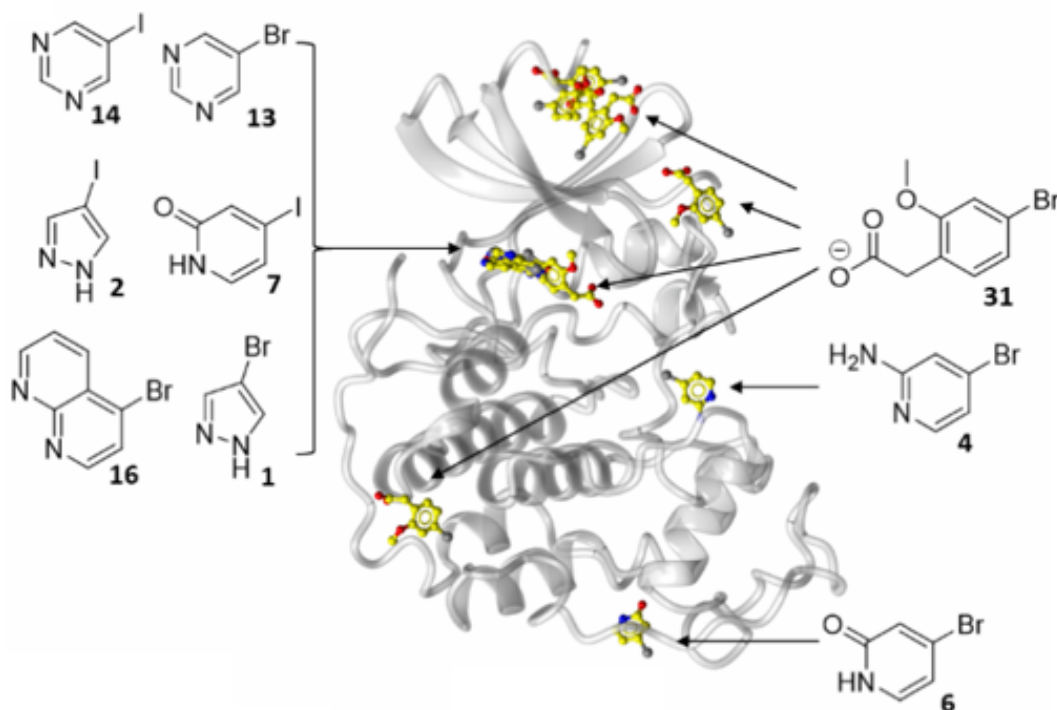


Figure 103; Overview of Fraglite binding events to CDK2 (white ribbon), depicting the sites identified by our Fraglite library (yellow)



Fragment **1** was found to bind in 2 different modes, in one case with the pyrazole-2-nitrogen accepting a hydrogen bond from the backbone of Leu83 and in the other case with the 2-nitrogen accepting a H-bond from Leu83 and the NH forming a H-bond with the backbone carbonyl of Glu81 (Figure 104).

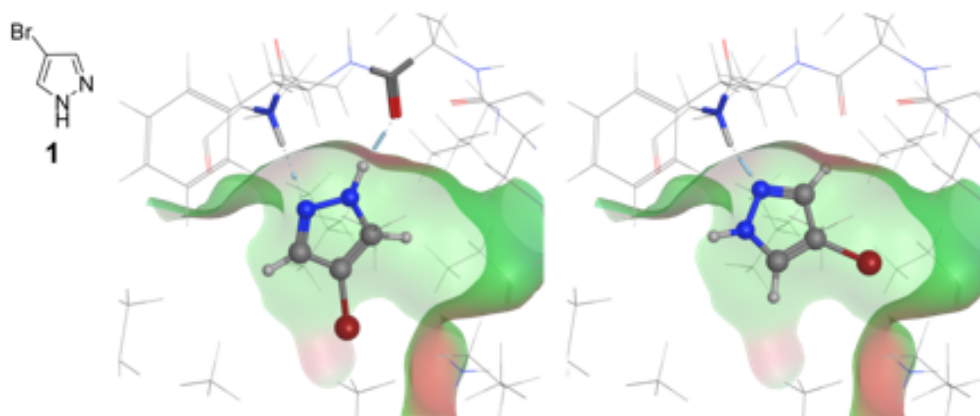


Figure 104; Interaction of fragment **1** with the ATP pocket

Fragments **2**, **7**, **13** and **14** showed very similar interactions (Figure 105-a) and fragment **16** bound to the ATP site through a halogen mediated interaction (Figure 105-b).

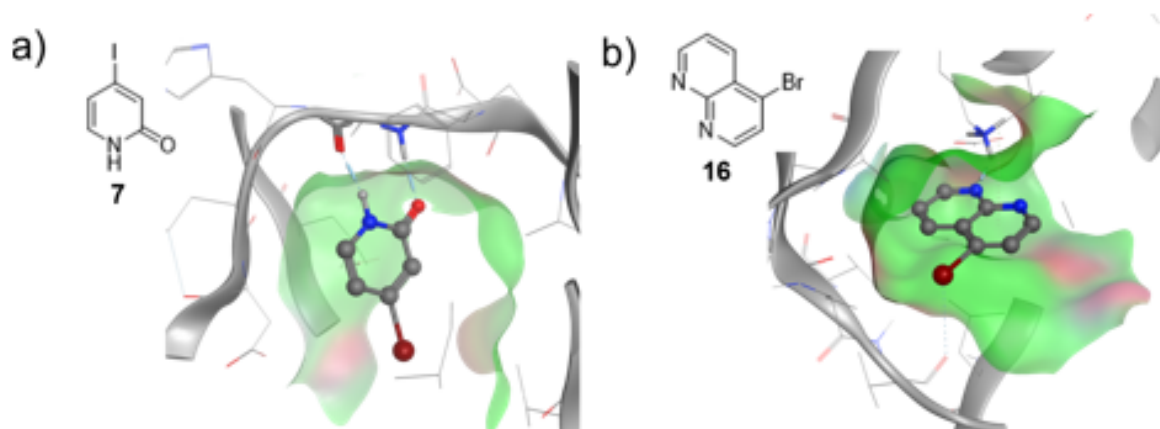


Figure 105; a) Fragment **7** forming donor-acceptor paired interaction with the NH and carbonyl of Leu83, b) Fragment **16** bound via halogen mediated interaction

It was also found that fragment **31** (2-methoxy-4-bromophenyl) acetic acid also bound to the hinge in an adjacent pocket, with the carboxylate interacting with the backbone NH and

sidechain hydroxyl of Thr14 and the amino group of Lys129, in addition a  $\pi$ -cation interaction with Lys33 was also observed (Figure 106).

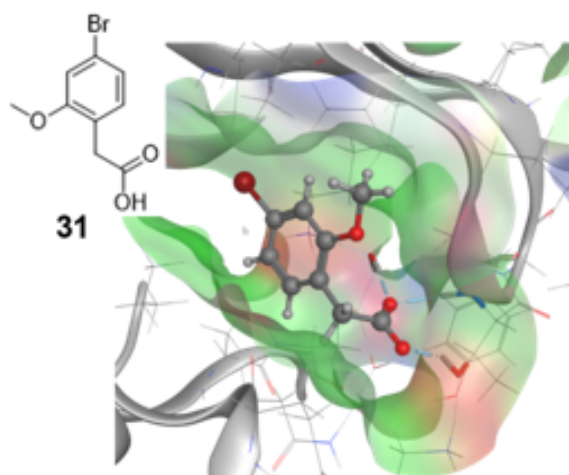


Figure 106, Fragment **31** bound to the hinge in an adjacent pocket

The hit rate obtained was surprisingly high and it was compared with Pan-Dataset Density Analysis (PanDDA) algorithm (state of the art statistical approach for detecting weak binding events in electron density maps<sup>172</sup>). PanDDA analysis was performed on 31 FragLite crystal datasets and could not identify the sites unveiled by FragLites **16** and **14** and several binding events found for FragLite **31**.

In addition to showing the location of binding sites, the Fraglites also identify the specific hydrogen bonding motif that makes the productive interaction, unambiguously identified due to the adjacent anomalous signal. This information can be used for optimization. The established 3-point hinge binding pharmacophore for CDK2 can be derived from the overlaid Fraglite signatures, and new molecules can be designed by merging adjacent or overlaid Fraglites (Figure 107).

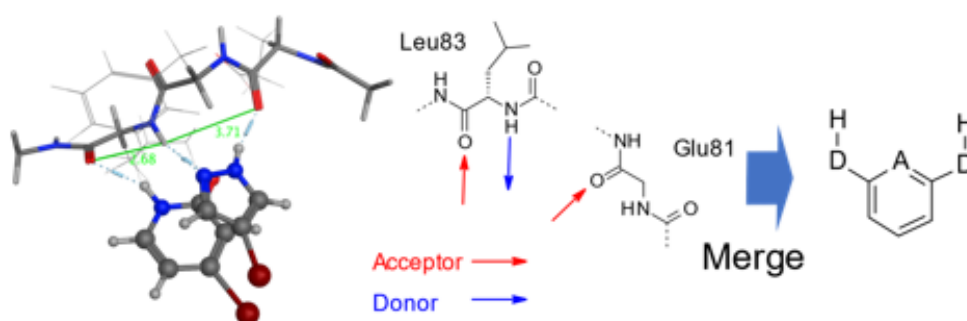


Figure 107; 3-point pharmacophore derived from 2 FragLites binding to the hinge



Molecules with the same binding motif were trialed to test their affinity with CDK2. Fragments **32** and **33** made the three suggested interactions by the binding motif proposed. Fragment **33** showed increased affinity by isothermal calorimetry (ITC) (Figure 108).

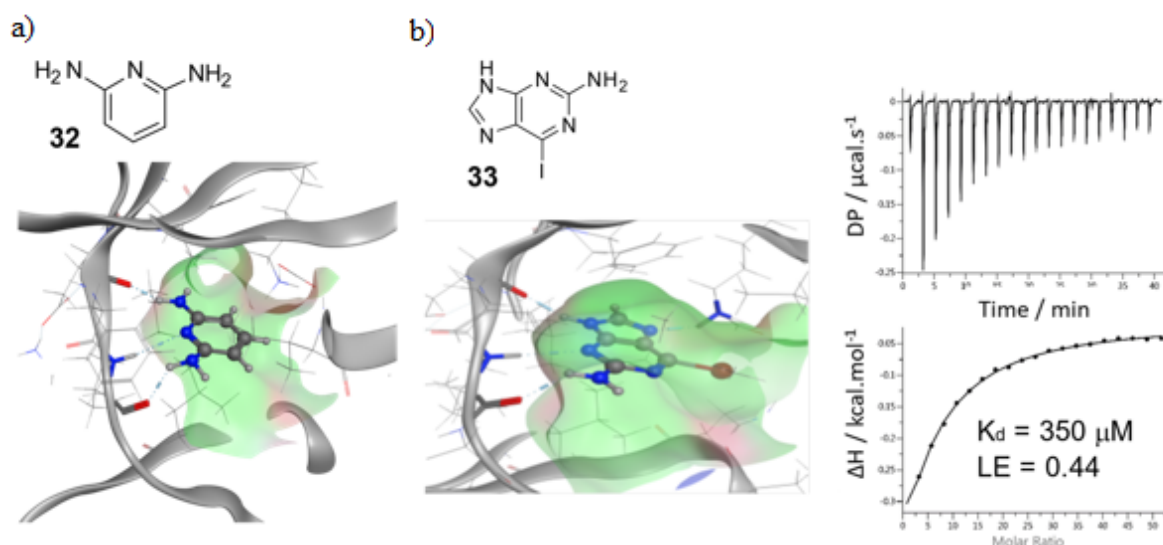
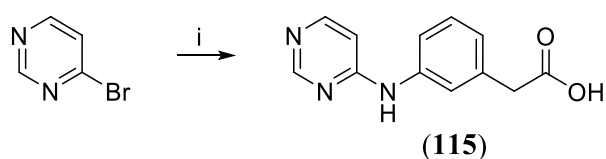


Figure 108; a) 2,6-diaminopyridine **32** fulfilling the 3-point pharmacophore and making the predicted interactions, b) 4-iodo-6-aminopurine **33** making the predicted interactions, showing binding with a measurable  $K_d$  by ITC

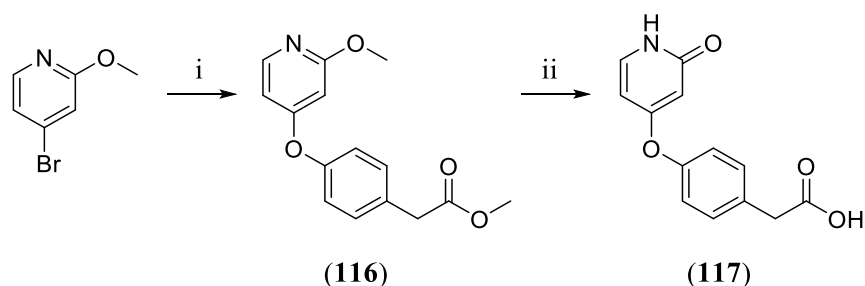
The orientation of fragment **31** binding in the ATP pocket gave us insights in how to grow the molecule. New molecules were designed with no additional screening. The structure of fragment **31** was explored in one case, tethering an aniline with a pyrimidine moiety, with two nitrogen atoms available as hydrogen bond acceptors (compound (**115**)) and in the other case tethering the aniline with pyridine, which has a combination of two binding hydrogen bonding groups (donors and acceptors) (compound (**117**)).

For the synthesis of compound (**115**) an  $S_NAr$  displacement of 4-bromopyrimidine with (3-aminophenyl)acetic acid was performed. The mixture stirred at  $80^\circ\text{C}$  for 3 hours affording the desired compound in 10% yield after purification (Scheme 38).



Scheme 38; i) (3-aminophenyl)acetic acid, DMF,  $80^\circ\text{C}$

In the case of compound **(117)** a copper catalysed etherification of 4-bromo-2-methoxypyridine with methyl (4-hydroxyphenyl)acetate was undertaken, the reaction was stirred for 24 hours at 120°C yielding the desired product **(116)** in 37% yield after purification. The intermediate was demethylated with 5 equivalents of pyridine hydrochloride at 180°C, which allowed the global demethylation, giving us compound **(117)** in 37% yield (Scheme 39).



Scheme 39; i) methyl (4-hydroxyphenyl)acetate (1.5 eq.), 3,4,7,8-tetramethylphenanthroline (cat.), CuI (cat.), Cs<sub>2</sub>CO<sub>3</sub>, toluene, 110° C, 24h; ii) pyridine hydrochloride (5 eq.), 180° C, 6h

The crystallization showed both compounds binding in a predictable way to the hinge region, with very similar interactions to those shown by fragments **(13)**, **(7)** and the carboxylate **(31)**. The overlay of these structures with the new compounds is shown in Figure 109.

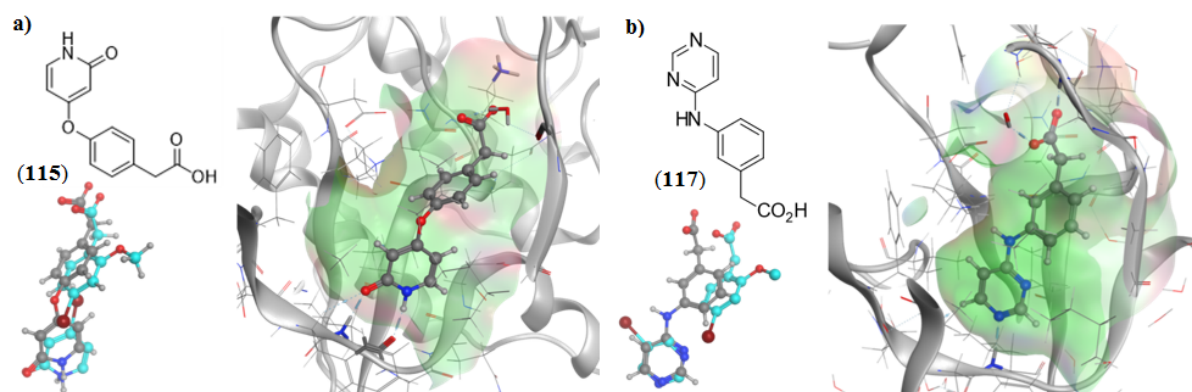
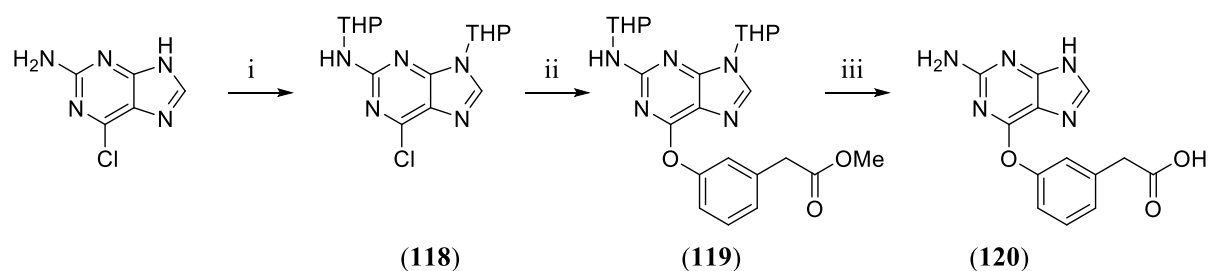


Figure 109; a) Overlay of **(115)** (grey carbon atoms), indicating orientation of the hinge motif, and **(31)** (cyan carbon atoms) as they bind to the CDK2 ATP pocket, b) Linked aminopyrimidine **(117)** binding to CDK2 and forming interactions with the hinge and Thr14, overlay shows the alignment of **(34)** (grey) with the original FragLites **(13)** and **(31)** (cyan)

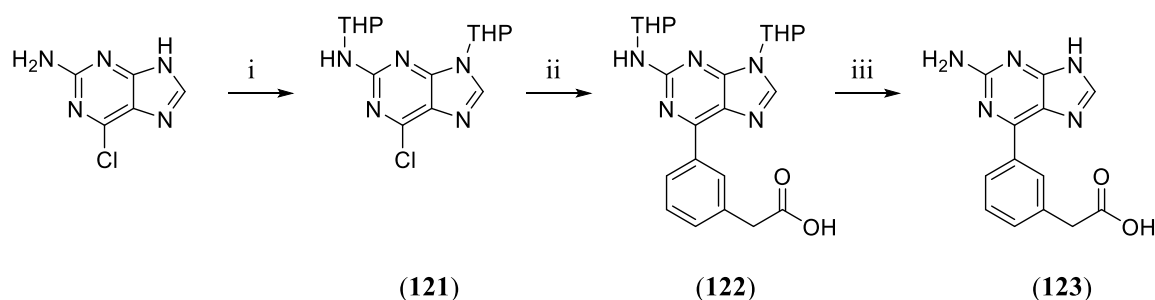
Fragment **(33)** was also expanded using the fraglite map, combining this fragment with carboxylate **(31)** affording two different biaryl compounds. For the synthesis of compound

(**120**), a standard method to achieve displacements at C-6 position of purines utilises a 6-chloropurine as starting material and proceeds via an intermediate with trialkyl ammonium as leaving group. The production of 6-trialkyl ammonium purine species is well described in the literature using DABCO.<sup>173</sup> Compound (**120**) was synthesized by S<sub>N</sub>Ar displacement after DABCO treatment of a DHP diprotected 6-chloro-9H-purine-2-amine with methyl 2-(3-hydroxyphenyl)acetate (Scheme 40).



Scheme 40; i) dihydropyran (10 eq.), aq. 2M HCl (cat.), DCM, r.t., 16h; ii) DMSO, DABCO (2 eq.), r.t., 3h; iii) 1) methyl 2-(3-hydroxyphenyl)acetate (1.33 eq.), DBU (1.1 eq.), DMF, 65° C, 72h; 2) aq. 2M HCl, THF, r.t., 4h

In the case of compound (**123**) (Scheme 41) a Suzuki-Miyaura coupling was performed directly in the DHP diprotected 6-chloro-9H-purine-2-amine and followed by acidic deprotection. The THP protecting group avoids protodeboronation during the cross-coupling (Scheme 41).



Scheme 41; i) dihydropyran (10 eq.), HCl (2M, cat.), DCM, r.t., 16h; ii) ethyl 2-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)acetate (2.5 eq.), K<sub>2</sub>CO<sub>3</sub> (2.5 eq.), Pd(PPh<sub>3</sub>)<sub>4</sub> (0.05 eq.), DMF:H<sub>2</sub>O (1:1), 140° C, 20', MW hv; iii) aq. 2M HCl, THF, r.t., 2h

The 2 biaryl compounds (**120**) and (**123**) were soaked in CDK2 crystals. Both compounds demonstrated concentration dependent binding by ITC (Figures 110 and 111).

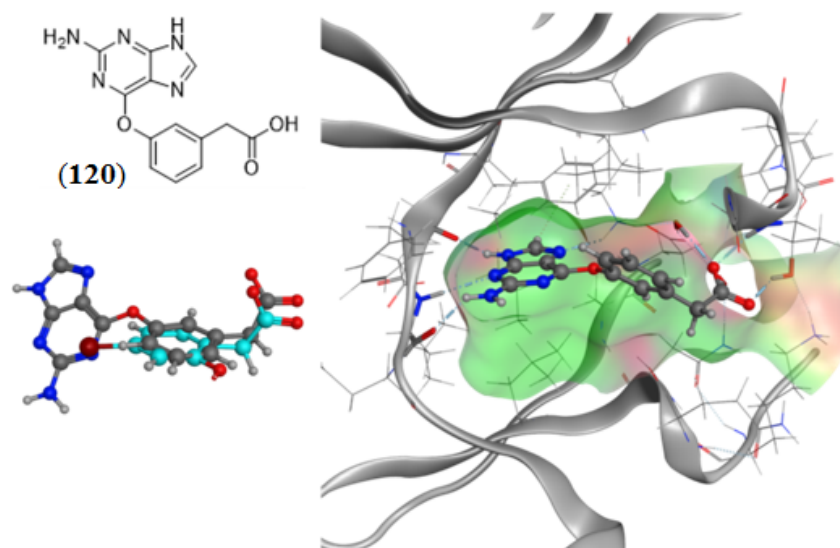


Figure 110; Linked pyridone biarylether (**120**) (grey) binding with the acid making comparable interactions to fragment **31** (cyan)

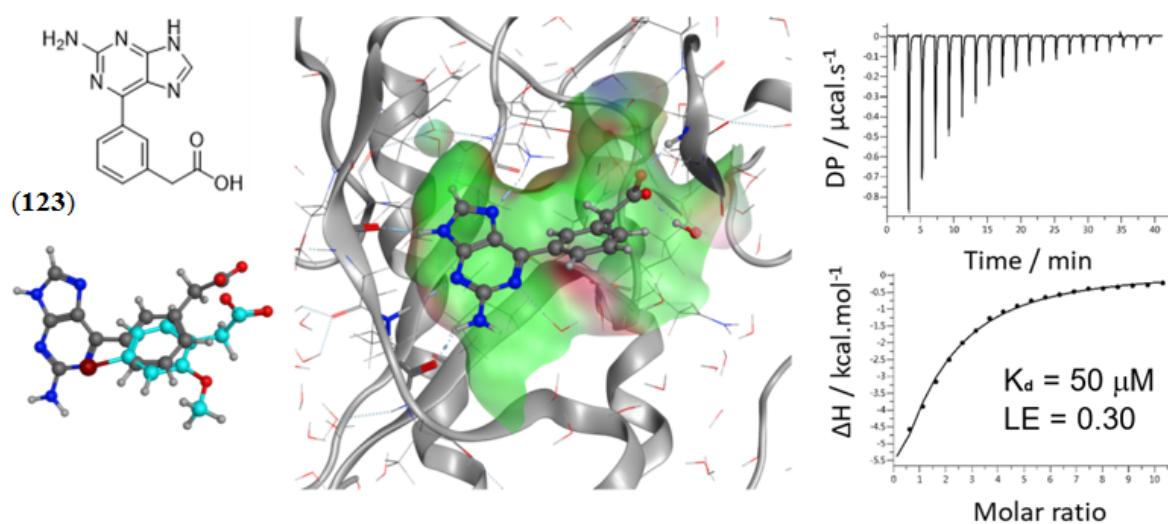


Figure 111; Linked biarylether (**123**) (grey) binding to CDK2 making comparable interactions to fragment **31** (cyan) and demonstrating concentration dependent binding by ITC

## Chapter 18: Conclusion and Future Work

Mapping the different binding sites of CDK2 with our Fraglite set was highly efficient. The orthosteric site was successfully identified as the principal binding site for several fragments. The different allosteric sites previously reported were also identified and a new one was also reported.

The efficacy of the Fraglite set for mapping CDK2 was high and with no need for a large and time-consuming screening campaign. The different crystal structures obtained during our investigation gave us key structural information about the location of the fragment binding and its orientation. These findings allowed us to design further analogues with combined structural motifs, which yielded a final compound (**123**) with concentration-dependent binding (measured by ITC;  $K_d = 50 \mu\text{M}$  and a  $LE = 0.3$ ). These results confirm the efficacy of the FragLite mapping approach to grow lead molecules shortening the screening process and without the synthesis of a large number of analogues.

For the future development of this strategy, new proteins should be explored in order to confirm the applicability of the FragLite mapping approach to other systems, and with the consequent optimization of a FragLite library. These findings will be shown in future publications with more challenging and less characterized proteins.

## Chapter 19: Introduction to MDM2 PROTACS for the androgen receptors as a prostate cancer treatment

### 19.1 PROTACs: Inducing protein degradation

#### 19.1.1 Background<sup>174</sup>

The classical small-molecule drug discovery approach relies on the occupation of a binding site in the target protein that produces a different structure. This change affects the protein function, and therefore a modification in the specific biological response of our target should occur (*i.e.* occupancy-driven strategy). At the same time, in order to achieve the desired biological response consistently, we have to consider the necessity of maintaining enough concentration of drug in our body to sustain target occupancy *in vivo*, which increases the risk of side effects.<sup>175</sup> An event-driven strategy to induce protein degradation is an alternative approach with different characteristics in which the target protein is tagged for elimination (Table 7).

Occupancy-driven strategy	Event-driven strategy
Stoichiometric activity	Sub-stoichiometric activity
Binding to active site required	Potential to target undruggable proteome
Site occupancy to block function	Restoration of function requires re-synthesis
Specificity defined by binding	Added layer of specificity

Table 7; Main differences of different medicinal chemistry approaches

One example of an event-driven strategy is the PROTACs (Proteolysis Targeting Chimeras) methodology but is not the only technique with event-driven characteristics. Other methods, like RNAi (RNA interference) can modulate the intracellular protein concentration, but very often engages off-target mRNA leading to side-effects.<sup>176,177,178</sup> We are also witnessing the dawn of CRISPR-Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats). This technology is able to directly modify the genome and find a way to a specific “gene knockout”, but its full potential is yet to be discovered.<sup>179, 180, 181</sup>

Novel approaches based on PROTACs take advantage of the cellular quality control systems to selectively degrade target proteins and have advantages over traditional strategies (occupancy-driven strategy). These advantages include the capacity to target a broad range of

proteins that traditionally have been considered undruggable.<sup>182,183</sup> Thanks to the genomic revolution many proteins have been identified to be directly associated to different types of cancer diseases,<sup>184,185,186</sup> but a lack of binding pockets for ligands which could potentially modulate the structure and the biological response associated to those<sup>187</sup> make this part of the proteome undruggable. With this regard, small-molecule-induced protein degradation can be considered a more interesting strategy. The hetero-bifunctional PROTACs consist of a ligand for the recruitment of E3 ligases, a linker, and a ligand to bind to the targeted protein, thus PROTACs have potential to eliminate “undruggable” protein targets, such as transcription factors and non-enzymatic proteins.

The approach is not limited to physiological substrates of the ubiquitin-proteasome system and, at the same time, inducing protein degradation can compensate for protein overexpression that often happens after loss of protein function due to inhibition. The consequent loss of therapeutic potential and pharmacological efficacy of our drug<sup>188</sup> is not a problem when the consistent degradation of the targeted protein takes place. Also, PROTACs can be advantageous to avoid resistance mechanisms that appear when druggable targets are recognized to have undruggable scaffolding roles.<sup>189, 190, 191, 192, 193, 194</sup> A very interesting example is EGFR (epithelial growth factor receptor), a receptor which independently of its kinase activity, maintains the basal intracellular glucose level, thereby preventing cells from undergoing autophagic death.<sup>195</sup> It has also been proved that the attachment of an E3 ligase ligand via a linker to a known inhibitor can add layers of selectivity to our warhead. The choice of linker and E3 ligase ligand are crucial for the formation of the ternary complex of our PROTAC with the target protein and the E3 ligase,<sup>196</sup> so degradation of the protein is dependent on the structure of the PROTAC. Also, recent evidence suggests that the action of PROTACs is catalytic, after degradation of one target protein the same PROTAC molecule can re-engage and degrade further proteins, which can reduce drug exposure and undesirable side effects.<sup>197</sup>

In summary, this new approach shifts the paradigm from inhibition of the protein to removing it by degradation. This gives us the opportunity of exploring a central challenge in biology, which is to understand, in a reliable way, the functions of a protein within the cell and the consequence of its removal. At the same time, the methodology allows us to control the protein expression levels without the need for genetic intervention.

### 19.1.2 Mechanism of action

The proteolysis targeting chimeras (PROTACs) are heterobifunctional molecules which recruit a specific protein target to an E3 ubiquitin ligase, resulting in the target's ubiquitination and degradation (also known as chemical knockdown). We can describe the way PROTACs work as the hijacking of the proteolysis machinery of the cell in which the target is tagged via poly-ubiquitin chains (for ubiquitin see Figure 112).

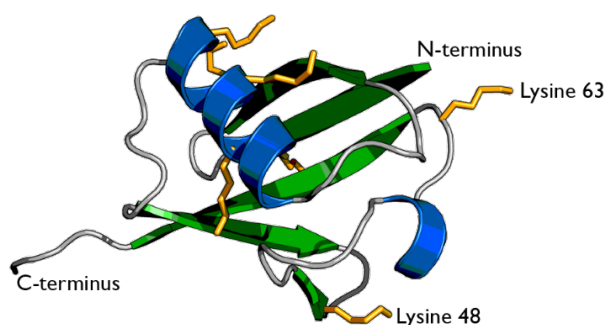
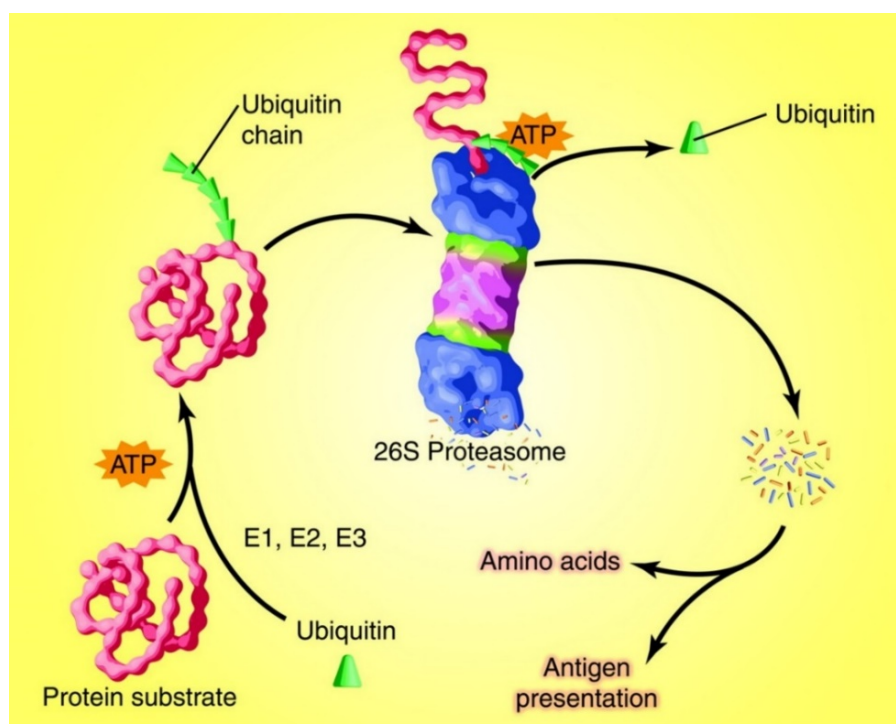


Figure 112; Ubiquitin is a regulatory protein that may be linked to other proteins; PDB ID: 1UBI

The “Ubiquitin-Proteasome Pathway” (UPS) consists on the conjugation of one or more ubiquitin molecules to other proteins for degradation by the 26S proteasome (Scheme 42).

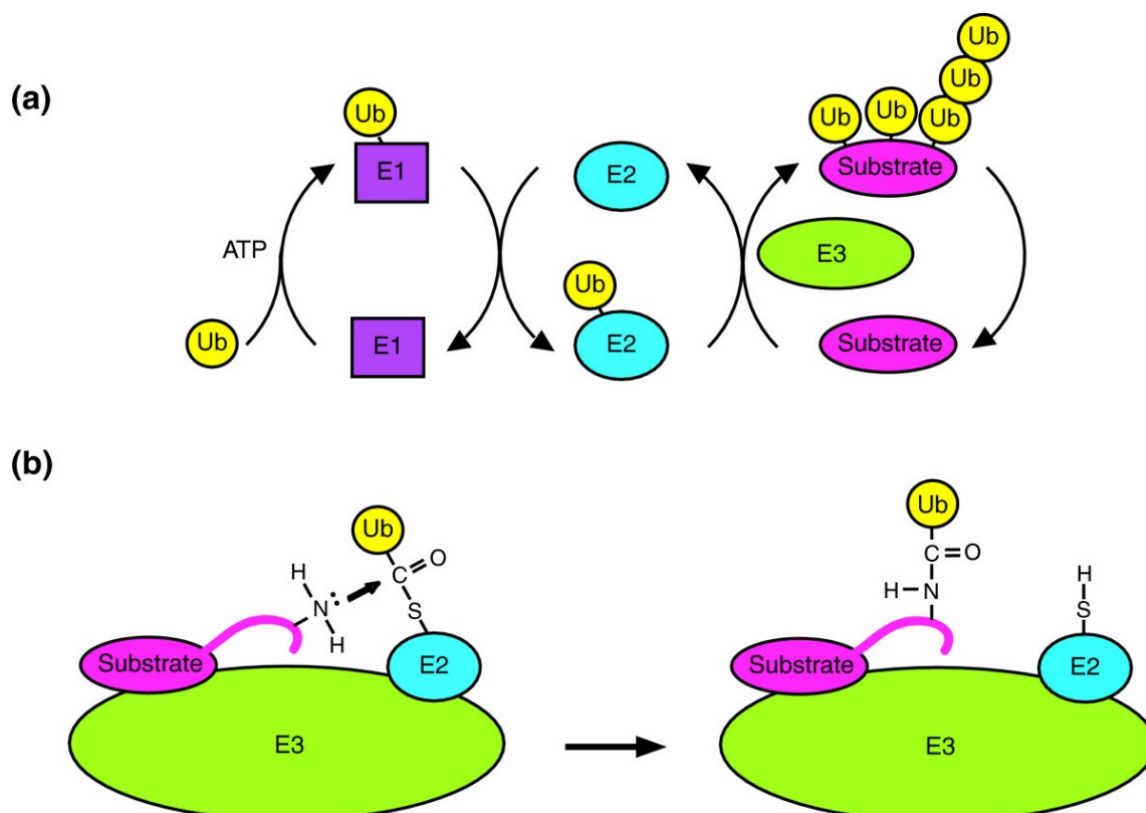


Scheme 42; Ubiquitin-Proteasome Pathway<sup>198</sup>



When discovered, the UPS was considered the first “natural” mechanism for the targeted destruction of misfolded proteins. The mechanism is based in a rationally designed enzymatic cascade that leads to the attachment of ubiquitin via covalent bonds to a lysine of the targeted protein.

The cascade (Scheme 43-a) is initiated by ubiquitin-activating enzyme (E1),<sup>199</sup> which forms a reactive thioester through an ATP-dependent process. The ubiquitin is then passed to ubiquitin-conjugating enzyme (E2),<sup>200</sup> again via formation of a reactive thioester. The last step of the cascade (Scheme 43-b) is performed by the ubiquitin ligase enzyme (E3),<sup>201</sup> which specifically recognizes a specific target protein and catalyses the transfer, induced by proximity, of ubiquitin from the E2 protein to the targeted protein via a lysine residue present in the surface of the protein to be tagged and subsequently degraded.



Scheme 43; **a)** Ubiquitin is first conjugated to the E1, then passed to the E2 protein; **b)** The E3 protein is responsible for substrate recognition, and transferral of the ubiquitin onto the substrate

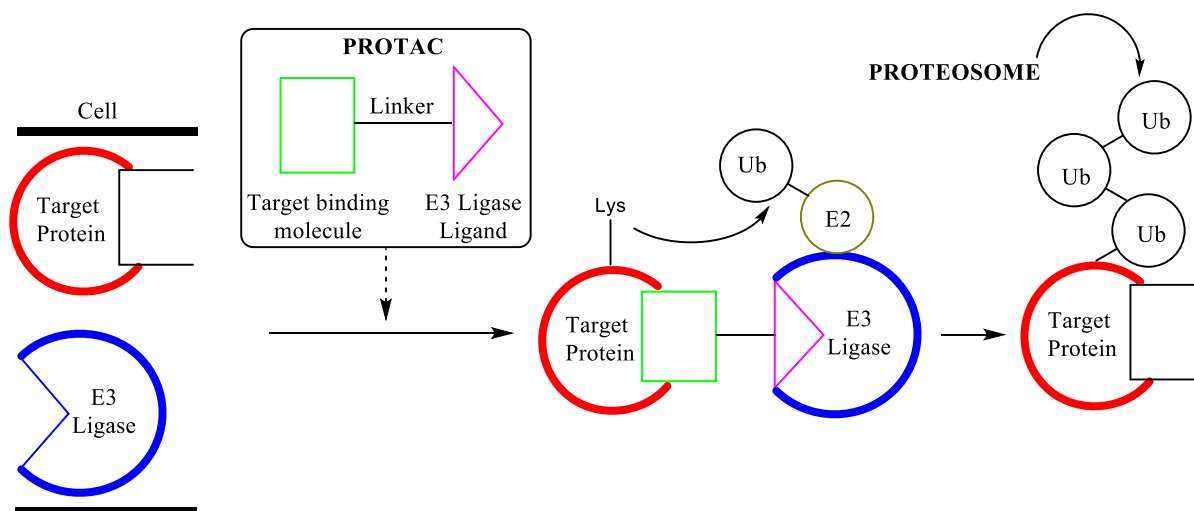
The specificity of the E3 ligase ligand for different proteins is very important for the degradation and management of protein levels in the cell. Several are usually considered for the design of successful PROTACs.

To regulate the concentration of proteins, the different E3 ligases recognize several specific binding motifs in the different targets. For example, the Von-Hippel-Lindau (VHL) E3 ligase binds to a core hydroxylated proline residue on hypoxia-inducible factor 1 $\alpha$  (HIF $\alpha$ )<sup>202</sup>, another example is the  $\beta$ -transducing repeat-containing protein ( $\beta$ TRCP) that binds specifically to phosphorylated residues on NF $\kappa$ B inhibitor- $\alpha$  (I $\kappa$ B $\alpha$ ) and  $\beta$ -catenin.<sup>203</sup> For the purpose of PROTACs a variety of E3 ligase ligands are commonly used (*e.g.*, cereblon ligand (CRBN), VHL ligand and  $\beta$ TRCP ligand). These act as substrate recognition subunits in multisubunit cullin-ring E3 ligase (CRL) complexes.<sup>204, 24</sup> Also Mouse Double Minute 2 homologue (MDM2) forms homo- or hetero-oligomeric E3 ligase complexes for productive ubiquitination.<sup>205, 206</sup>

As ubiquitin has seven lysines, substrate-linked ubiquitin can be further modified to produce unique polyubiquitin chains, which are recognized by the 26S proteasome, which will induce degradation of the tagged protein.

Summarizing, the PROTACs concept could be described with 3 main aspects (Scheme 44):

- Targeting a ubiquitin ligase to recognise a protein of your choice
- Bringing the protein you are interested in into close proximity to an ubiquitin ligase
- A bivalent PROTAC ligand is used to reversibly bind the target to the E3 ligase and promote ubiquitination of the target protein



Scheme 44; Main steps of the PROTACs strategy

## 19.2 Alternative approaches to protein degradation

### 19.2.1 HSP90 inhibitors

Induced protein degradation is not a completely new concept. The inhibitors of the molecular chaperone heat shock protein 90 (HSP90) entered clinical trials in 1999. The drug with the ability to bind HSP90 and subsequently degrading client proteins was under big expectations due to the discovery of upregulation of the chaperone as a consequence of interaction with cancer cells to drive proliferation and survival pathways.<sup>207</sup> All the different drugs designed target the ATP-binding domain<sup>208,209, 210</sup> leading to the degradation of HSP90 client proteins.<sup>211,212</sup> The antiproliferative effects of these compounds is related to destabilization of the Raf-1 protein, one of the client proteins of HSP90, resulting in inhibition of MAPK. After more than 30 clinical trials and a wide exploration associated to HSP90 binders, none has achieved approval from the FDA. Then main reason is its poor pharmacokinetic profile, mainly associated to the hepatotoxic benzoquinone moieties present in the ligands (Figure 113).<sup>213, 214, 215</sup> On top of that HSP90 target proteins associate with more than half of the human kinome and attempts to achieve selective degradation have been extremely challenging.<sup>216</sup>

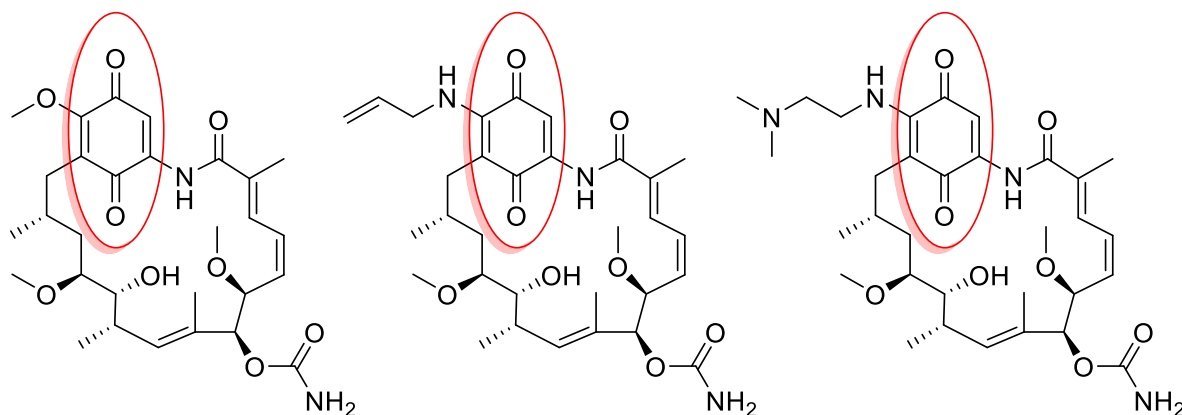


Figure 113; **Geldanamycin** (left hand-side) and its derivatives

### 19.2.2 Selective hormone receptor degraders (SERDs & SARDs)

The estrogen receptors (ER) are nuclear hormone receptors that are suitable for modulation with small-molecule drugs. After binding to the hormones (estrogens), the ER dimerises and translocates from the cytosol to the nucleus where it binds to DNA to activate transcription. The proteins expressed after estrogen signalling have a great influence on the cardiovascular, musculoskeletal, immune and reproductive systems in males and females. It has also been

proved that ER signalling has a clear contribution to breast cancer and its progression, and this type of carcinomas are responsive to hormone deprivation therapy (Figure 114).

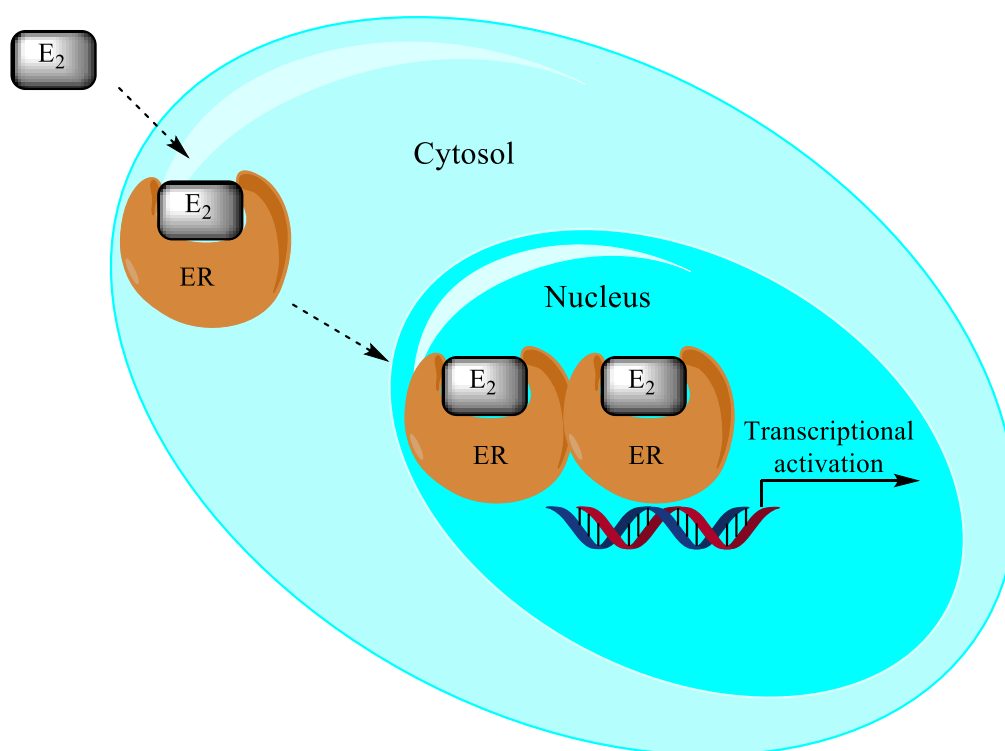


Figure 114; Activation and translocation of the estrogen receptor (ER) to the nucleus to activate transcription ( $E_2$  = oestradiol)

Depending on several factors, selective estrogen receptor modulators (SERMs) act as antagonists or agonists, being an antagonist in breast tissue but an agonist in the uterus, enhancing in the last case the risk of endometrial cancers.<sup>217</sup> Tamoxifen is a SERM,<sup>218</sup> and this drug is used to treat breast cancer but with a small increased risk in the development of uterine cancer. On top of that, a long term estrogen deprivation could lead to pharmacological resistance due to compensatory feedback signalling and/or mutations in the estrogen receptor, which impair the pharmacological efficacy of tamoxifen making necessary the intervention of a new treatment.

Following the strategy of induced degradation via destabilization of the protein, fulvestrant modifies the structure of the ER creating a larger hydrophobic surface due to the presence of an alkylsulfinyl moiety (Figure 115, thus activating the degradation of the receptor<sup>219</sup> by hydrophobic tagging.

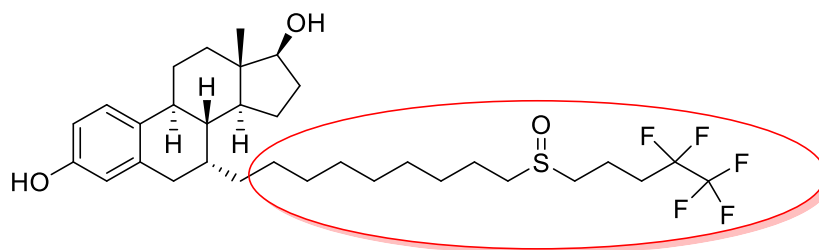


Figure 115; **Fulvestrant** with alkylsulfinyl moiety highlighted

Fulvestrant has lower risk of endometrial cancer<sup>220</sup> and, despite several pharmacokinetic issues, is the only FDA-approved selective estrogen receptor degrader (SERD) for ER-positive metastatic breast carcinomas after development of resistance to tamoxifen. Newer generations of fulvestrant are under investigation.<sup>221</sup> The limited success of this drug was still encouraging, and the same strategy was investigated with androgen receptors (ARs) in prostate cancer, to develop selective androgen receptor degraders (SARDs).

As a nuclear hormone receptor, the AR responds to mainly two hormones, testosterone and 5 $\alpha$ -dihydrotestosterone (5 $\alpha$ -DHT). For this reason prostate cancer is dependant on AR signalling, thus antiandrogen therapy with flutamide or bicalutamide has demonstrated good results.<sup>222</sup> Again, resistance is developed, leading to castration resistant prostate cancer (CRPC)<sup>223</sup> via overexpression or amplification of the ARs.<sup>224</sup> Much work is being undertaken and there are several clinical trials ongoing (including PROTACs from Arvinas<sup>225</sup>), but at the moment no known SARDs are available in the market.

As a conclusion, inducing ER and AR degradation is a very interesting approach that remains active. Targeting the degradation of either receptor not only fights the compensatory upregulation associated to inhibition of a function but could also disrupt resistance mechanisms. Moreover, the transformation of an antagonist into an agonist via mutation of the receptor should not be a problem for an induced degradation strategy. At the same time, high systemic exposure associated to classic drugs could be reduced due to the catalytic nature of the PROTACs.

### 19.2.3 Hydrophobic tagging (HyT)

This strategy consists of a heterobifunctional molecule, one end able to be bound to the protein of interest (POI) and the other end able to bind the cell quality control mechanism through hydrophobic moieties,<sup>226</sup> which induce protein destabilization. One proposal for the



An alternative hydrophobic tagging method was discovered using adamantane (Figure 118).<sup>228</sup> This approach is also based on a bifunctional molecule, one end engages the cell quality control mechanism via a hydrophobic moiety, and the other end binds to the protein of interest.<sup>229</sup>

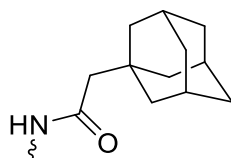


Figure 118; **Adamantane** Moiety

Recently the hydrophobic technology (HyT) has been expanded to ARs. A high affinity agonist<sup>230</sup> was bound to an adamantyl moiety, making a degrader from an agonist<sup>231</sup> (Figure 119).

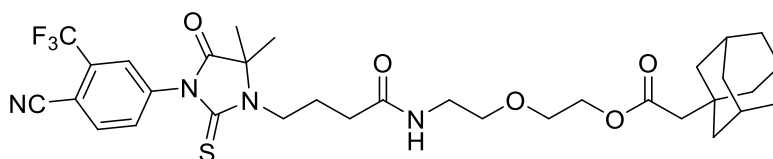


Figure 119; **SARD279**

The HyT shows the scope and possibilities of modular post-translational protein knockdown by targeting proteins traditionally considered undruggable. The main weakness of this method is that the hydrophobic moieties can potentially bind to human serum proteins having as a result the loss of free drug available to induce degradation. Furthermore, it is still not clear if the degradation is induced by the hydrophobic moiety (Boc<sub>3</sub>Arg or Adamantane) or by the induced destabilization of the protein itself leading to self-degradation without proteasome intervention, however both of these approaches would work. By contrast, PROTACs that induce proximity-induced ubiquitynation of a targeted protein via E3 ligase ligands are already in several clinical trials and showing more promising *in vivo* performances and better pharmacokinetic profiles.

#### 19.2.4 Peptide-based PROTACs

The first peptide-based PROTAC was reported in 2001.<sup>232</sup> A bifunctional molecule with the ability to harness the ubiquitin-proteasome system (UPS) by recruiting an E3 ligase to the

protein of interest, and leading to ubiquitination, and consequently degradation of the protein was described (Figure 120).

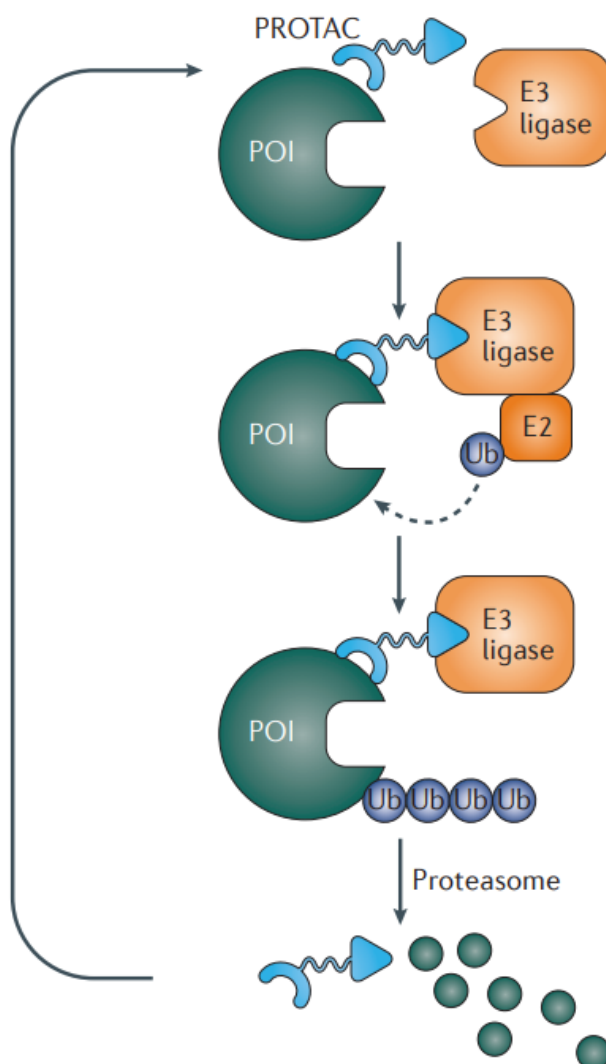


Figure 120; Representation of a peptide-based PROTAC strategy

The peptide-based PROTACs have at one end a phosphopeptide (E3 ligase ligand) portion of NF- $\kappa$ B inhibitor- $\alpha$ (I $\kappa$ B $\alpha$ ). This peptide recruits the F-box protein  $\beta$ -transducin repeat-containing protein ( $\beta$ TRCP), a component of the S-phase kinase-associated protein 1 (SKP1)–cullin 1–F-box E3 ligase complex (SCF $^{\beta$ TRCP). On the other end, in a similar way to other PROTACs, the enzyme inhibitor ovalicin (protein ligand), covalently binds to the target protein, methionine aminopeptidase 2 (MetAP2).<sup>233</sup>

The potencies of these PROTACs were not high enough (micromolar range) and cellular permeability was extremely low; moreover, the dependency on I $\kappa$ B $\alpha$  phosphorylation for E3 ligase recruitment was limited by the action of cellular phosphatases, therefore improvements



were needed. The crucial one was the switch to a smaller peptidic E3 ligase ligand fragment in which phosphorylation is not required for functioning. The E3 ligase ligand Von Hippel-Lindau (VHL) (Figure 121) recognizes the E3 ligase hypoxia-inducible factor 1 $\alpha$  (HIF1 $\alpha$ ).<sup>234, 235</sup>

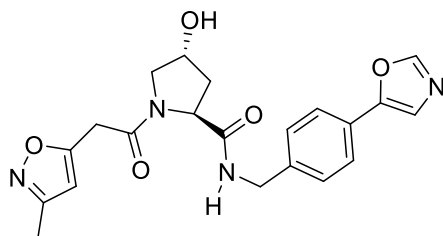


Figure 121; 1<sup>st</sup> generation **VHL** ligand<sup>236</sup>

With this new E3 ligase ligand in hand a new PROTACs era started. As a first example, a PROTAC targeting AR comprised DHT linked to the HIF1 $\alpha$  peptide fragment<sup>237</sup> was developed and proved to lead to AR depletion in HEK293 cells (Figure 122-a). Analogous PROTACs for ER degradation using ER inhibitors were positively tested, reducing proliferation of breast cancer cell lines<sup>238</sup> *in vitro* (Figure 122-b).

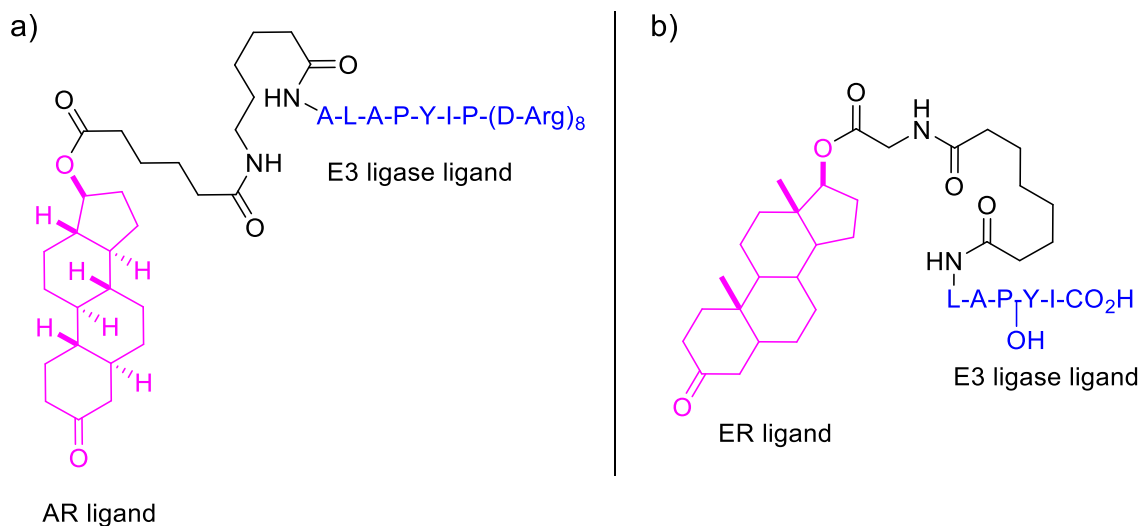


Figure 122; **a)** Peptide-based PROTAC for AR, **b)** Peptide-based PROTAC for ER

However, the potencies of peptidic PROTACs are very low, remaining usually in the micromolar range in the best cases and showing pharmacokinetic issues like poor cell penetration. Having these problems in mind a new type of PROTACs was to be developed.

### 19.2.5 All Small-Molecule PROTACs

Considering the cellular problems of peptide-based PROTACs and its limitations an important effort was done focused on the generation of all small-molecule PROTACs.

In 2008 the synthesis of a PROTAC based on Nutlin-3a (Figure 123) to recruit E3 ligase mouse double minute 2 (MDM2)<sup>239</sup> to target the AR<sup>240</sup> showed an induced AR degradation of 10  $\mu$ M concentration in cells. The first “all small molecule” made the bifunctional strategy much more drug-like, especially from a pharmacokinetic point of view. These results had implications for the potential study and treatment of various cancer types with increased androgen receptor levels.

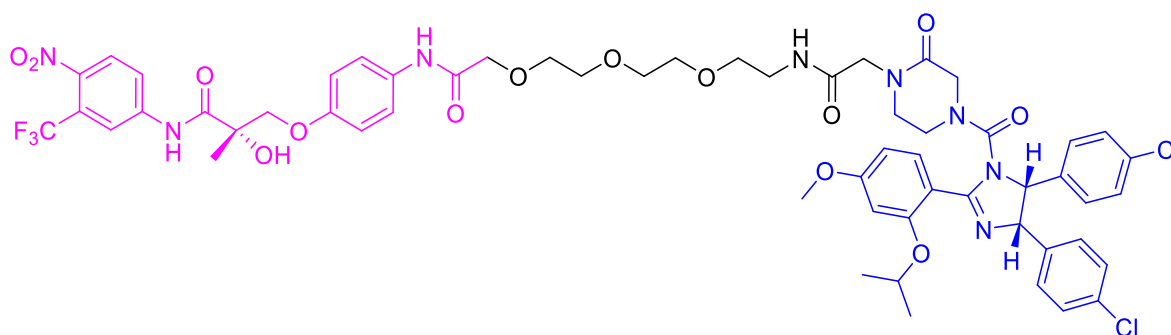


Figure 123; **Flutamide** (AR ligand) in pink, nutlin-3a (MDM2 ligand) in blue

The potency was lower than with the peptidic VHL ligand but newer generations of E3 ligase ligands have been developed with increased potency<sup>241</sup> and improved pharmacokinetic profiles.

### 19.2.6 The case of SNIPERs

Based on the same approach for protein degradation, another type of small molecule-based PROTAC called SNIPER (specific and non-genetic inhibitor of apoptosis proteins (IAPs)-dependent protein eraser) was reported by the Hashimoto group in 2010. Using bestatin (Figure 124) ester to recruit the E3 ligase cellular IAP1 (cIAP1)<sup>242</sup> the degradation of cellular retinoic acid binding protein I (CRABPI) and CRABPII was achieved. However, bestatin was not specific and these esters induced the polyubiquitination and subsequent degradation of the recruited E3 ligase cIAP1 reducing the efficiency of this technology and showing the first example of a PROTAC degrading the E3 ligase meant to be recruited.

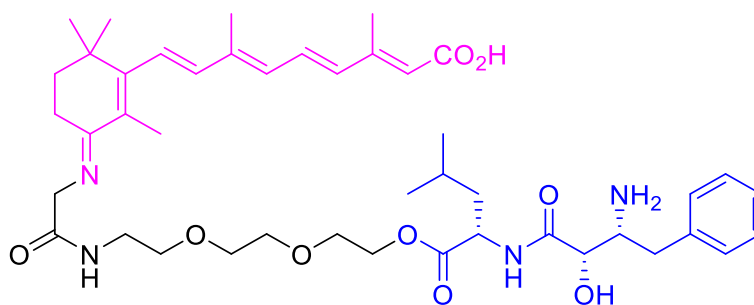
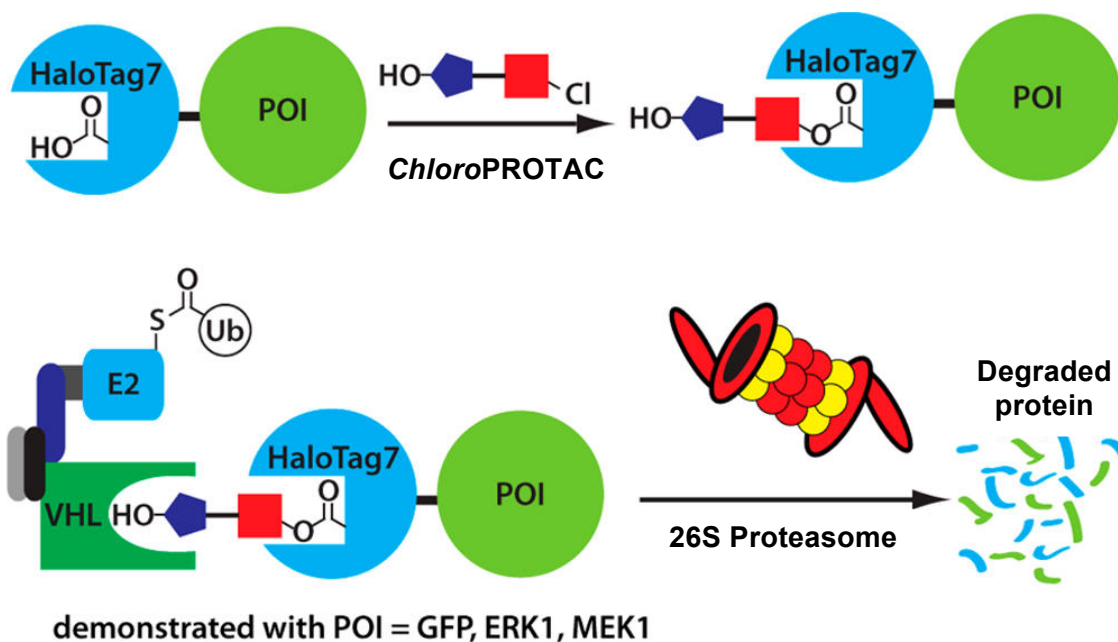


Figure 124; CRABPII inhibitor (protein ligand) in pink, bestatin (cIAP ligand) in blue

### 19.2.7 Halo-PROTACs

To overcome some of the pharmacokinetic handicaps related to PROTACs, a new idea was proposed. Introducing a label in the protein of interest with a *HaloTag* and then using *ChloroPROTAC* that can react with the *HaloTag* creates a covalent bond into the cell. With this method we can avoid the use of poorly permeable compounds and build our PROTAC inside the cell (Scheme 45).<sup>243</sup>



Scheme 45; (POI = protein of interest; VHL (Von Hippel Lindau protein) = E3 ligase ligand)

### 19.2.8 New directions identifying a replacement for the HIF-1 $\alpha$

A replacement for the HIF1 $\alpha$  peptide fragment to provide an alternative E3 ligase for the all-small-molecule PROTAC approach was identified for the VHL E3 ligase in 2012 ( $K_d=5.4$



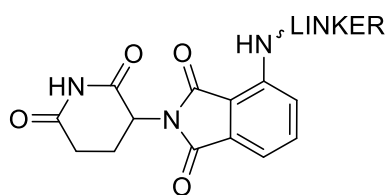


Figure 126; **Pomalidomide** (CRBN ligand)

#### 19.2.10 The catalytic nature of PROTACs

In 2015 two “all small” PROTACs were reported. In one case using Cereblon as an innovative E3 ligase ligand for BET domain protein Brd4<sup>249</sup> (Figure 127) and in the other using VHL.<sup>250</sup> These molecules demonstrated the catalytic nature of PROTACs *in vivo*, which was subject of controversy.

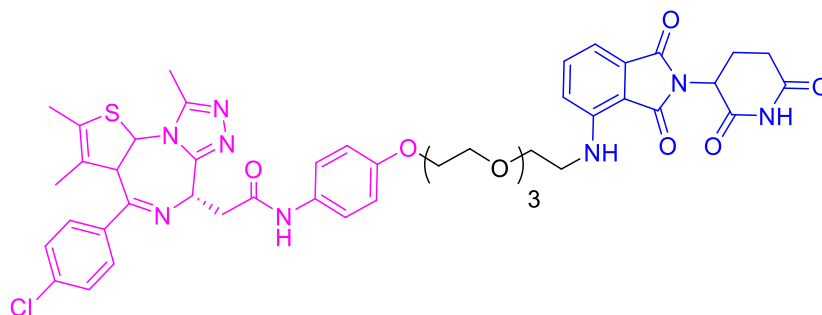


Figure 127; Pomalidomide (CRBN-based E3 ligase ligand) in blue and Brd4 inhibitor in pink

It has been recently confirmed by a comprehensive real-time degradation and recovery profiles for targets quantifying degradation rates, maximal levels of degradation ( $D_{\max}$ ) and time frame at  $D_{\max}$ . These studies also showed that the ternary complex stability has a direct relation with the capacity of the PROTAC for degradation.<sup>251</sup>

#### 19.2.11 Added layers of selectivity

In 2015, Alessio Ciulli reported a study where the induced degradation strategy provides selectivity towards a specific bromodomain. The Bromo- and Extra-Terminal (BET) proteins BRD2, BRD3, and BRD4 play master roles in the transcription regulation and epigenetics and are associated to the development of cancer diseases. The pan-BET selective bromodomain inhibitor JQ1 lacks selectivity and limits the scope of the inhibitor. By tethering JQ1 to a ligand for the E3 ubiquitin ligase VHL, the PROTAC was unexpectedly selective to BRD4.<sup>252</sup>

### 19.2.12 First crystal structure of the 3<sup>ary</sup> complex

In 2016 Alessio Ciulli published the first crystal structure of the key ternary complex ligase-PROTAC-target species (Figure 128). Its structural elucidation remained elusive until then and can be a key for future structure-based-design PROTACs. The results elucidate how PROTAC-induced *de novo* contacts dictate preferential recruitment of a target protein into a stable and cooperative complex with an E3 ligase for selective degradation.<sup>253</sup>

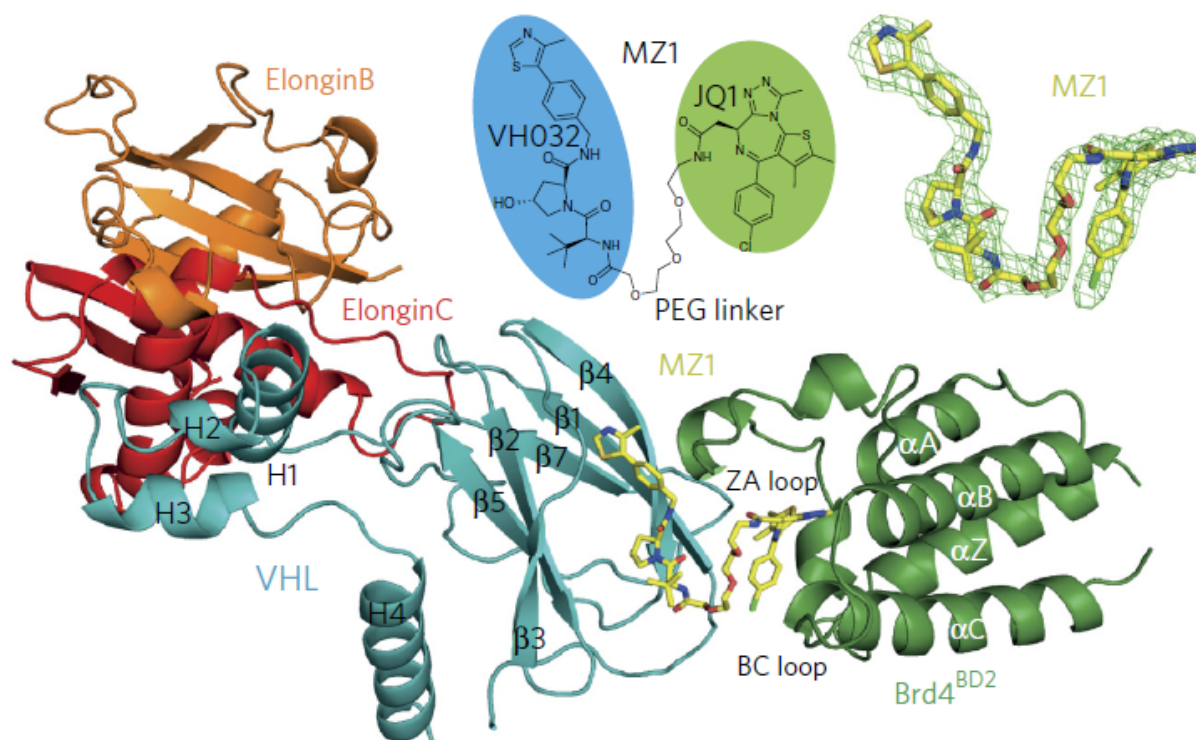


Figure 128; Crystal structure of the 3<sup>ary</sup> complex

#### 19.2.1 The importance of the linker

The importance of the linker has been highlighted as a major contributor of the different conformations that the PROTAC can adopt in space. The influence of the linker, and consequently, the flexibility and shape of the molecule produced has an influence in the activity of the PROTAC binding the two targets of interest (protein and E3 ligase) and creating the ternary complex.<sup>254</sup>

In 2017 Craig Crews published a study with a rational identification of a degrader targeting a serine/threonine TANK-binding kinase-1.<sup>255</sup> This article states that the progression to a potent PROTAC have generalized key structural elements associated with degradation activities which can be rationalized.

A new study in 2019 demonstrated that with the right linker tethering position and the appropriate linker length several E3 ligase ligands can be used successfully for degradation.<sup>256</sup>

### **19.2.2 Challenges and opportunities ahead**

PROTACs technology is having a great impact in the development of molecules able to induce degradation. To date, this approach has been mostly used to target druggable proteins, however the ability to degrade transcription factors and scaffolding proteins would increase the current target drug space and give access to new therapeutic strategies.

However, there are potential problems with PROTACs. As we can clearly see, this approach will always use molecules with very high molecular weights, also these are likely to have many Hydrogen Bond Donors and Acceptors (HBD and HBA) and a high Polar Surface Area (PSA). As a consequence, the cell penetration is challenging but the potency of these compounds might be increased by the catalytic nature of ligands.

### **19.2.3 From initial concept to industry**

The Craig Crews research group has been focused in controlled proteostasis. In 2001 Crews developed, in collaboration with Ray Deshaies, the concept of PROTACs. 12 years later, in 2013, Crews founded Arvinas. The company, which main business is the development of PROTACs to treat cancer and other diseases, partnered with Merck and Genentech in 2015 to develop PROTAC-based drugs. In 2018 Arvinas signed a \$830M plus deal with Pfizer, the pact centers on the discovery and development of PROTACs across multiple disease areas. The promise of induced protein degradation is very attractive for industry<sup>257</sup> (Figure 129) and has led to the creation of companies focused on this technology (Arvinas was launched in 2013<sup>258</sup>, and C4 therapeutics was launched in 2016<sup>259</sup>).

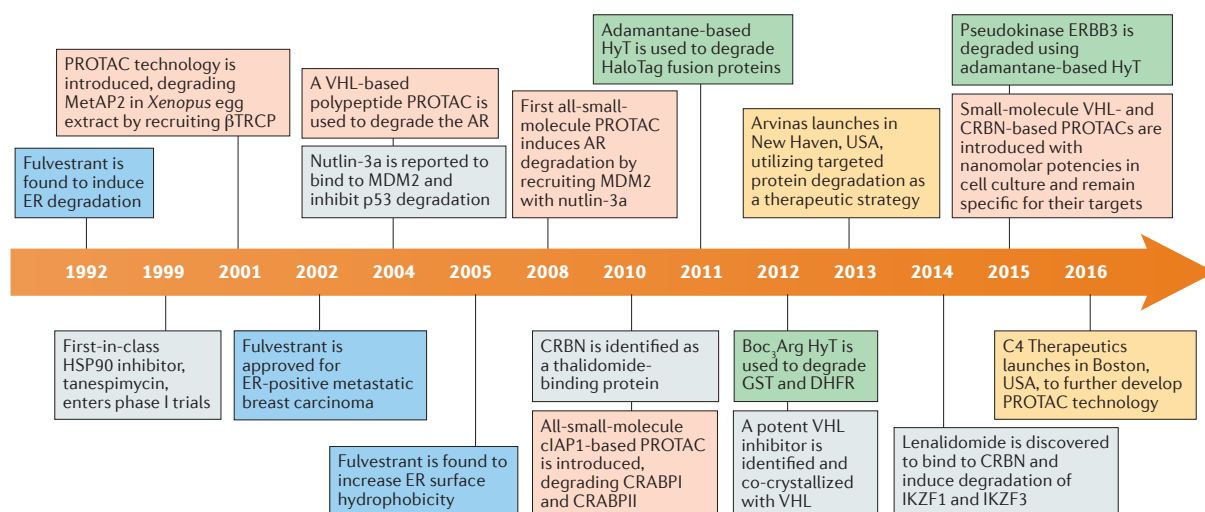


Figure 129; Timeline of protein-induced degradation. Events related to **estrogen receptors (ER)** in blue, **PROTACs technologies** in red, **hydrophobic tagging (HyT)** in green, **companies** in yellow

Recently, Arvinas has announced that a PROTAC-based drug candidate had advanced into phase I clinical trial in 2017<sup>260</sup> for mCRPC. Currently there are no drugs in the market based on PROTACs. The success of these clinical trials can pave the way for a new generation of therapeutic drugs based in a completely new approach.

#### 19.2.4 Recently reported MDM2 induced degradation via PROTACs as a viable treatment for leukaemia

PROTACs is a very recent and promising technology and important papers have been reported while this thesis was in preparation. In a recent publication by Shaomeng Wang *et al.*<sup>261</sup> was reported an MDM2 induced degradation strategy based in PROTACs that achieves complete and durable tumour regression *in vivo* and was reported as a new class of anticancer agent. This publication is a nice example of the successful development of a PROTAC, and it will be summarized here to illustrate the potential of this strategy.

In the studies reported here, an MDM2 inhibitor (**MI-1061**) was used as a small molecule ligand. After studying the co-crystal structure (Figure 130) the solvent exposed area of the ligand, with a carboxylic acid on the phenyl ring, was proposed as the tethering site for the design of the bivalent compound. **MI-1061** was tethered with different linkers to two different E3 ligase ligands for Cereblon: thalidomide and its derivative lenalidomide (Table 8 and 9).



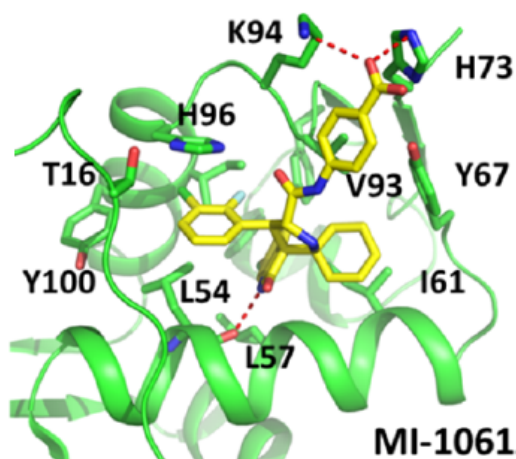
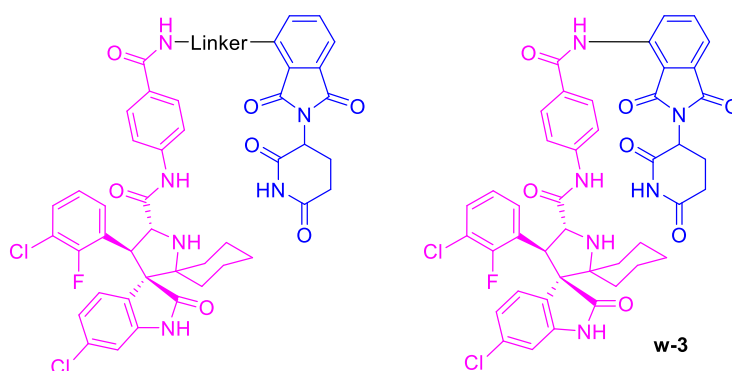


Figure 130; Modelled structure of **MI-1061** in complex with MDM2

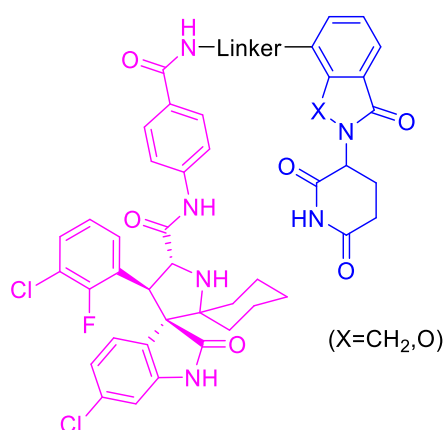
In the case of thalidomide (Table 8) a compound where the MDM2 inhibitor and the E3 ligase ligand are directly linked to examine the importance of the linker (Table 8-entry w-3) was investigated. The analysis shows that compound w-3 failed to achieve acceptable levels of inhibition compared to other compounds. Changing the linker to different lengths and the introduction of ethylene glycols instead of alkyl chains (w-9, w-10) has an impact in the inhibition of cell growth.



Compound No.	Linker	IC <sub>50</sub> (nM) cell growth inhibition
w-3	Void (no linker)	68.4 ± 36
w-8		100 ± 27
w-9		8 ± 2
w-10		5 ± 1

Table 8; IC<sub>50</sub> values after linker exploration

It is also known that modifications on the Cereblon binding portion of the degrader molecule can significantly change the levels of induced degradation of the protein. Exchanging thalidomide for lenalidomide (Table 9) gave new compounds (w-11, w-12) with improved IC<sub>50</sub> values. The rigidification and shortening of the linker was also investigated with compounds w-14 and w-15, obtaining with the former the most potent compound (Table 9-entry w-14).



Compound No.	Linker	X	IC <sub>50</sub> (nM) cell growth inhibition
w-11		CO	29 ± 8
w-12		CH <sub>2</sub>	5 ± 5
w-14		CH <sub>2</sub>	1.5 ± 0.5
w-15		CH <sub>2</sub>	3.9 ± 0.4

Table 9; MDM2 Degradors with Modifications of the tethering to the CRBN Ligand Portion

The analysis of the co-crystal structure of CRBN E3 ubiquitin ligase complexed with lenalidomide shows the possibility of tethering the linker in 3 different positions (Figure 131), this is very important for the development of PROTACS as the formation and stability of the 3<sup>ary</sup> complex can vary substantially, and the different possibilities must be explored. However, in this case no improvement was observed.

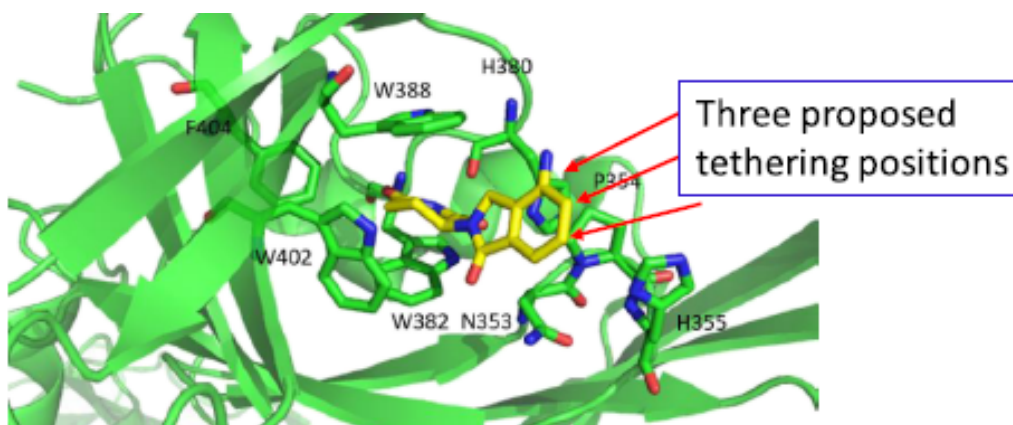


Figure 131; PDB ID: 4CI2

Different E3 ligases can target the same protein, and the specificity of those for a specific target is hard to predict and may vary depending on each PROTAC's design. For this reason several E3 ligase ligands are usually trialled in a given project. In addition to CRBN/Cullin 4A E3 ligase complex, the von Hippel-Lindau protein (VHL-1)/cullin 2 E3 ligase system was also explored. The synthesis of several analogues gave compounds with good potencies but considerably inferior to the ones previously synthesized series. Therefore, after an analysis of all the data collected, the most promising compounds **w-13** (MD-222) and **w-14** (MD-224) were selected for further evaluation.

To confirm degradation of MDM2 and subsequent accumulation of p53 in a dose dependant manner, Western blotting analysis were conducted in leukaemia cell lines. Both (**w-13**) MD-222 and (**w-14**) MD-224 induced degradation in different cell lines, with effective depletion of MDM2 (green box) and accumulation of p53 (blue box) (Figure 132).

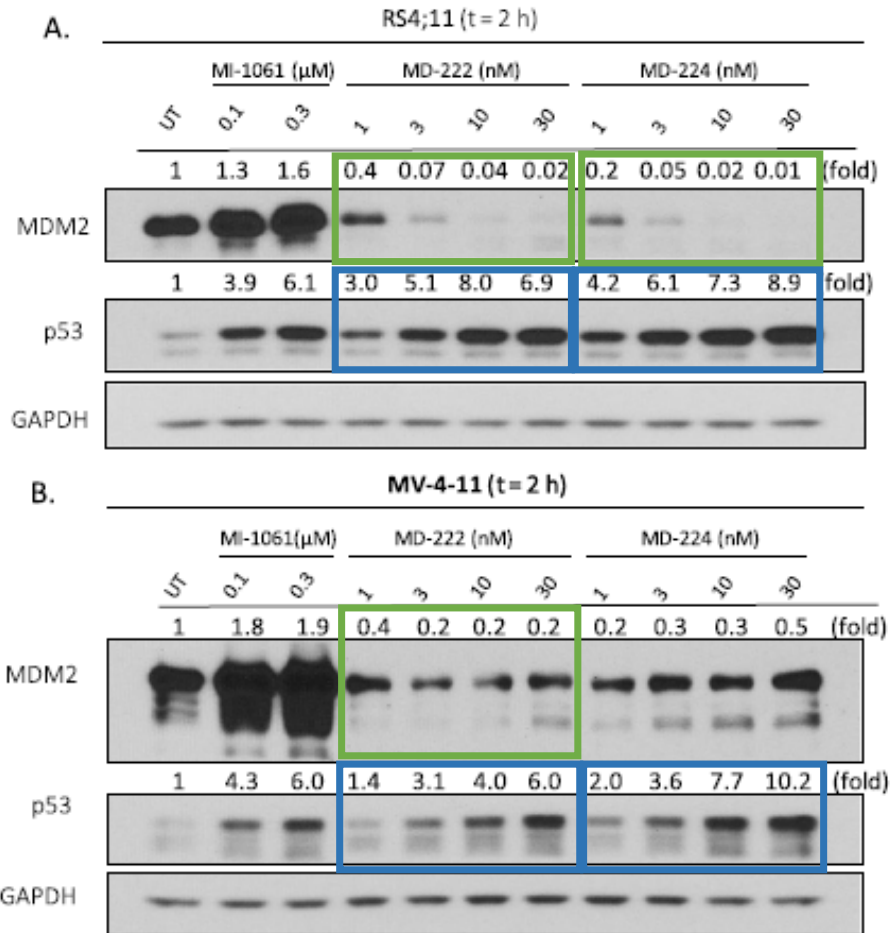


Figure 132; Western blotting analysis for MDM2 (green box) and p53 levels (blue box) after treatment with MDM2 inhibitors MI-1061 and the MDM2 degraders MD-222 and MD-224, (RS4;11 cell line in fig. 129-A and MV4;11 cell line in fig. 129-B)

One experiment generally used to confirm degradation upon addition of PROTACs is based on blocking the CRBN/Cullin 4 E3 ligase complex upon addition of an excess of lenalidomide. Western blotting analysis shows that lenalidomide actually blocks MDM2 degradation (green box) and accumulation of p53 protein (blue box) is induced despite the presence of PROTACs (Figure 133).

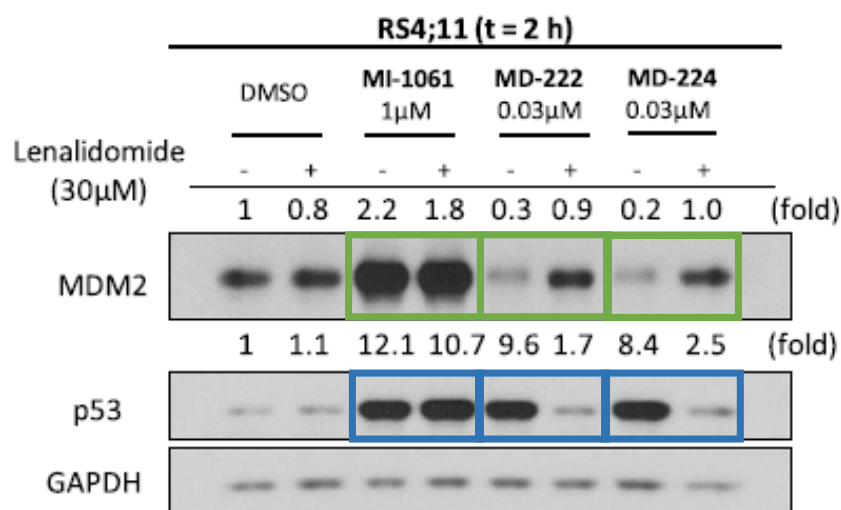


Figure 133; Western blotting analysis of both MDM2 degraders in presence of lenalidomide excess. MDM2 degradation in the presence of PROTACs (green box) and accumulation of p53 (blue box)

We can also verify the PROTACs activity with the addition of a proteasome inhibitor. In the case that the cell degradation machinery is impaired the PROTACs cannot induce degradation and only MDM2 inhibition patterns should be observed. After preincubation with proteasome inhibitors MG-132 and PR-171 or a neddylation inhibitor, MDM2 degradation is inhibited in a dose dependant manner (Figure 134-green box).

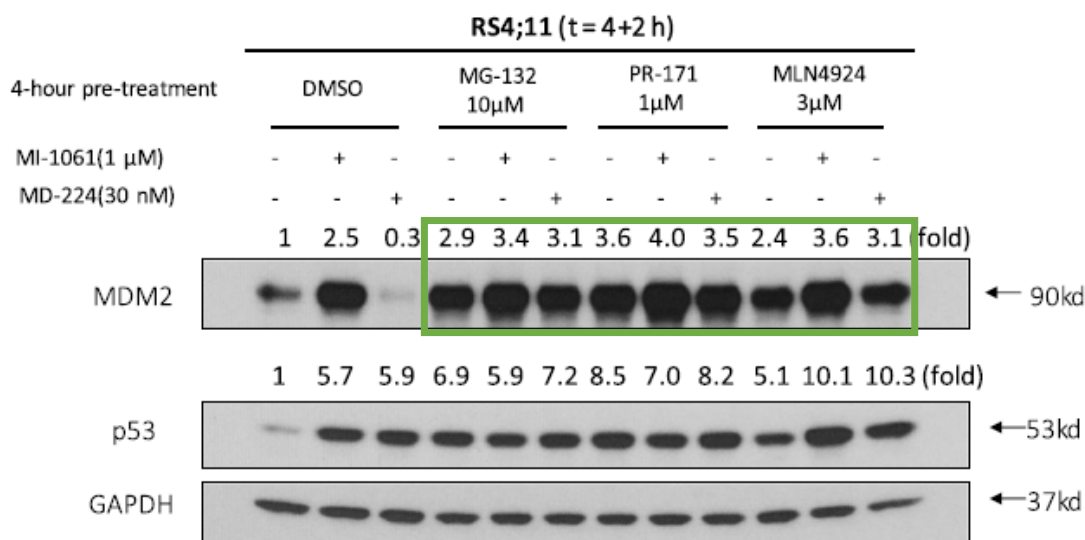


Figure 134; Western blotting analysis of MDM2 (green box) and p53 proteins in RS4;11 cells

After several analysis in different leukemia cell lines and p53 (wild type and mutant), and having very promising pharmacodynamics data, MD-224 was tested *in vivo* in comparison with MI-1061 MDM2 inhibitor (Figure 135).

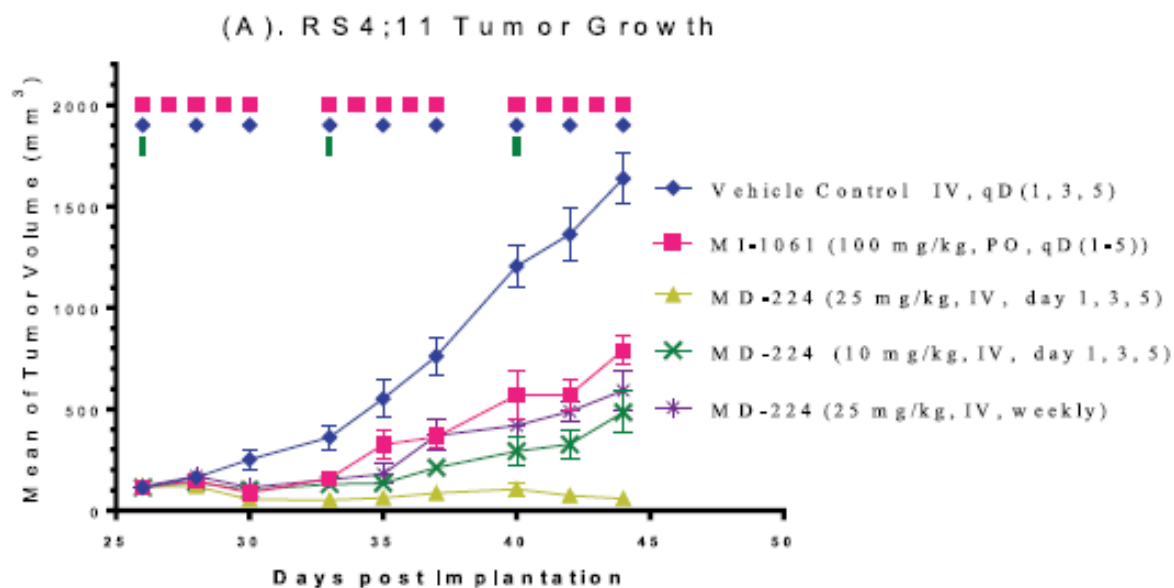


Figure 135; *In vivo* antitumor activity of **MI-1061** and **MD-224** in the RS4;11 xenograft model in mice

The MDM2 inhibitor effectively retards tumour growth, but MD-224 can achieve more potent activity with significant tumour volume loss and no signs of toxicity or weight loss.

MD-224 effectively induces inhibition of cell growth and is >10-100 times more potent than the MDM2 inhibitor used as positive control (MI-1061). This compound is highly efficacious and will be further evaluated as a potential new therapy for the treatment of human leukaemia.

### 19.3 Androgen Receptors and PROTACs

#### 19.3.1 The case of androgen receptor (AR) and flutamide

The Androgen receptors (AR) is a member of the nuclear receptor superfamily and they are transcription factors ligand-dependant. ARs are activated by binding of endogenous androgens (mainly DHT), regulate several processes in the male reproductive system. They play a significant role in the development of prostate cancer and castration resistant prostate cancer (CRPC). Mechanisms of resistance can emerge to antagonists of the AR (like flutamide, enzalutamide, bicalutamide or apalutamide (Figure 136)) by mutations that convert these antagonists into agonists.<sup>262</sup> A mutation of the AR is believed to be responsible for metastatic prostate cancer (mPCa) cells ability to survive through mechanisms which rely on retaining the AR signalling pathway.

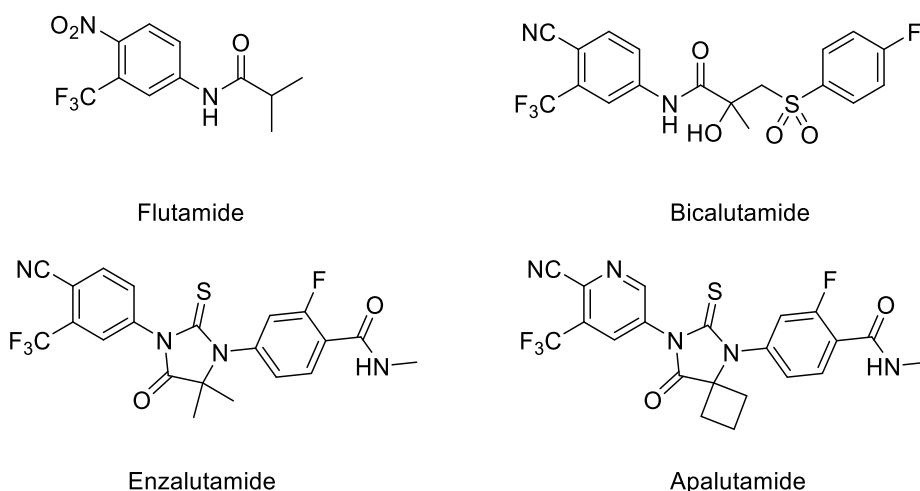
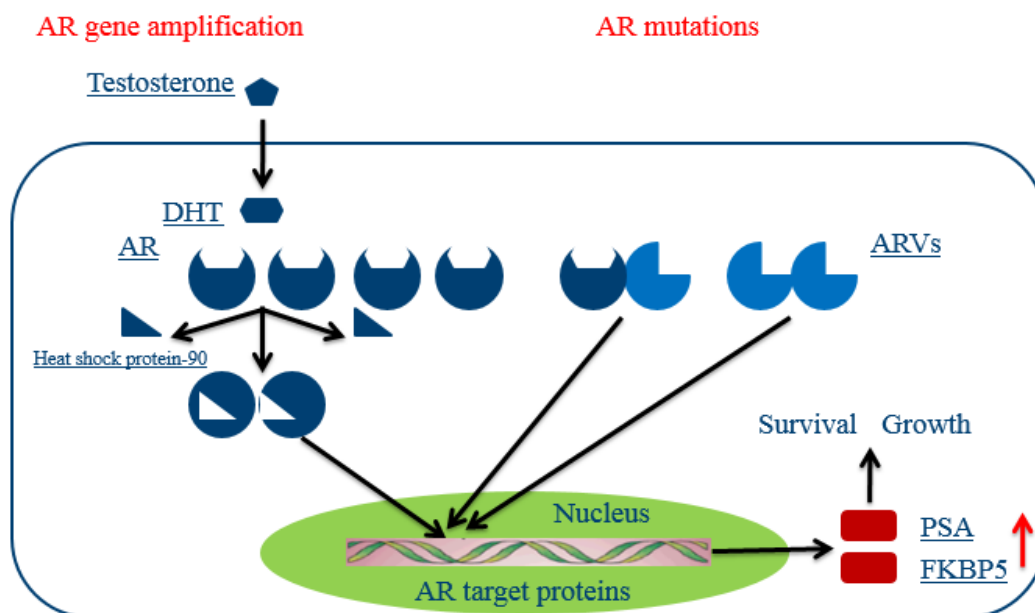


Figure 136; Different Androgen Receptor inhibitors (ARi)

There are several known androgen receptor variants (ARVs), ARF876L is driven by enzalutamide and ARW741L/C and ARH874Y/ART877A are known to be driven by flutamide. These ARVs along with AR amplification provide prostate cancer cells a mechanism for Androgen Deprivation Therapy (ADT) evasion (Scheme 46) and currently, the reduction of available androgen/testosterone/dihydrotestosterone (DHT) by chemical or surgical means is the only way to fight prostate cancer.



Scheme 46; Evasion strategy through Androgen receptor variants (ARVs) of ADT

Intensive studies have been carried to produce analogues of the available antagonists (flutamide (Figure 137),<sup>263</sup> bicalutamide, enzalutamide, apalutamide, etc...) to overcome

resistance. Also innovative lines of research are under development like nicotinamide derivatives as a novel type of nonsteroidal AR antagonists, or curcumin natural pigment analogues, which have shown antiandrogenic activity.

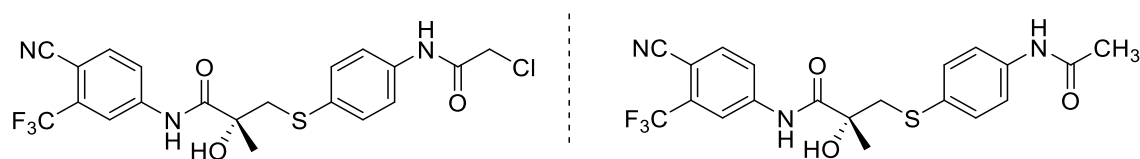
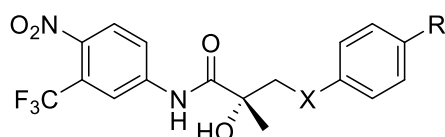


Figure 137; Previous analogues of flutamide

A series of nonsteroidal ligands based on flutamide were synthesized by Marhefka *et al.*<sup>264</sup> as 2<sup>nd</sup> generation ligands for the AR. These ligands were designed initially to eliminate metabolic sites with fast metabolism and low activity. Thanks to this research, more potent compounds with improved pharmacokinetic properties were discovered, by changing the sulfur-linked compounds to oxygen-linked compounds (Table 10).



	X	R	K <sub>i</sub> (nM)	Relative Activity
DHT	-	-	0.27	100
R-2	S	NHC(O)CH <sub>3</sub>	4.90 ± 0.20	50
S-5	O	F	6.11 ± 0.19	43.40 ± 2.60
R-5	O	F	225 ± 15	N.D.
<b>S-6</b>	O	NHC(O)CH <sub>3</sub>	3.98 ± 0.70	92.92 ± 7.00
S-7	O	C(O)CH <sub>3</sub>	36.6 ± 2.4	9.67 ± 1.50
S-8	O	C(O)CH <sub>2</sub> CH <sub>3</sub>	6.07 ± 0.14	75.41 ± 11.50
S-9	O	Cl	9.56 ± 0.66	64.01 ± 11.20
S-10	O	Br	11.64 ± 0.36	17.02 ± 2.90
S-11	O	I	29.96 ± 2.69	16.00 ± 2.90
S-12	O	CH <sub>3</sub>	34.77 ± 3.50	8.95 ± 2.50
S-13	O	OCH <sub>3</sub>	13.69 ± 0.68	17.40 ± 3.30
S-14	O	NHC(O)OC(CH <sub>3</sub> ) <sub>3</sub>	336 ± 71	4.43 ± 1.30
S-17	NH	F	7.96 ± 0.43	21.33 ± 2.22

Table 10; K<sub>i</sub> and *in vitro* activity of 2<sup>nd</sup> generation nonsteroidal AR ligands (Marhefka *et al.*)



## 19.4 MDM2 as an E3 ligase ligand

### 19.4.1 MDM2 biology

MDM2 is a protein and an E3 ligase that has several roles inside cells. One of its most studied functions is to control the levels of the transcription factor p53, which is achieved by MDM2 forming an auto regulatory feedback loop by virtue of the ability of p53 protein to activate the *MDM2 gene* and therefore MDM2 protein expression. MDM2 is a 491 amino acid protein that contains a deep hydrophobic binding site. For binding between MDM2 and p53, a triad of amino acids (Phe19, Trp23 and Leu26) in p53 binds to a shallow cleft in MDM2. Binding between p53 and MDM2 is driven essentially by hydrophobicity, with multiple Van der Waals interactions formed (Figure 138).<sup>265</sup>

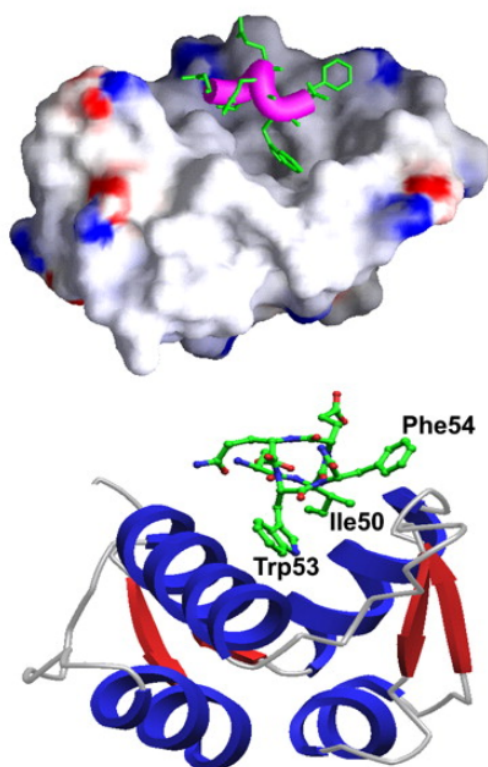


Figure 138; MDM2 bound with the p53 turn II (49 –54) peptide. The residues of p53 peptide are represented in a ball and stick mode with nitrogen atoms in blue and oxygen atoms in red<sup>266</sup>

MDM2 is a RING E3 ligase, so it can be used as a target for a PROTACs strategy. As we have described, the active site of MDM2 has been studied by identifying the key interactions between p53 and MDM2, but also importantly, recent studies have shown that there is a 30 aminoacid acidic region of MDM2 that is critical for binding to the E2 ligase and promote ubiquitination.<sup>267</sup> Several inhibitors for MDM2 have been developed with thorough structural characterization of its interactions, these inhibitors are available for PROTACs development.

### 19.4.2 MDM2 inhibitors

The first group of small molecules developed to target the E3 ligase MDM2 are called Nutlins. These ligands can displace p53 from MDM2 with  $IC_{50}$  values in the range of 100-300 nM. Through crystal structure determination it was found that the Nutlin compounds successfully bind to the p53 pocket in MDM2 with its cis-imidazoline structure allowing a halophenyl group to sit in the Trp23 and Leu26 pocket (Figure 139).

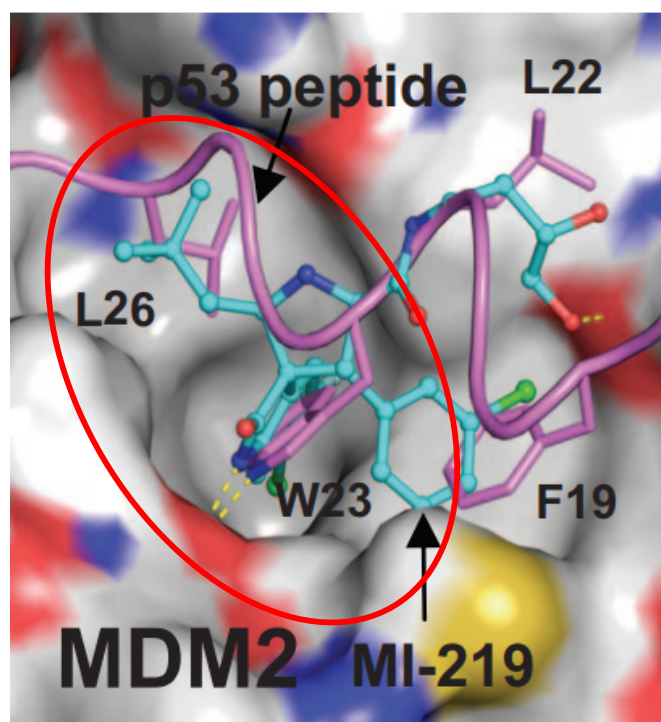


Figure 139; MI-219 (Nutlin) in complex with MDM2 and superpositioned to the MDM2– p53 crystal structure. MI-219 is shown in ball and stick representation<sup>268</sup>

This cis-imidazoline (Figure 140) provided a strong scaffold for further exploration of the molecule and to direct binding towards the MDM2 binding site.<sup>269</sup>

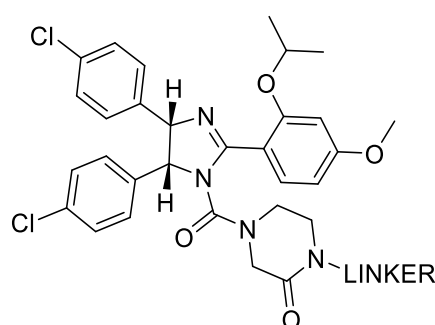


Figure 140; **Nutlin-3** compound used for the synthesis of PROTACs<sup>270</sup>

## Chapter 20: Project Rationale

The main objective of this project is to synthesize innovative PROTACs for the androgen receptors (ARs) that could be used as chemical probes for future anti-cancer treatments focused on prostate cancer.

### 20.1 Design of a hetero-functional molecule

As we already know a PROTAC is divided in three parts and the design will be rationalized for each one of these parts:

#### 20.1.1 E3 ligase ligand

MDM2 is an E3 ligase and is overexpressed in around 20% of cancer cells. The advent of potent and specific ligands for MDM2 offers the opportunity to develop an optimised MDM2-based PROTAC. A considerable number of small molecule inhibitors of this specific protein-protein interaction have been reported and successfully entered clinical trials.<sup>271, 272, 273, 274</sup>

The participation of several research groups and pharmaceutical companies have provided us with a broad spectrum of inhibitors of which a pyrazolopyrrolidine scaffold reported by the Novartis was selected to be used as the E3 ligase ligand.

The design and characterization done to achieve this new inhibitor is remarkable and based on previous experience of Novartis for this particular interaction.<sup>275, 276, 277</sup> The initial design relies on the occupancy by appropriate moieties of three essential pockets of the N-terminal domain of MDM2 which will recognize the residues of the transactivation domain of p53 (Phe19, Trp23 and Leu26).<sup>278</sup> For a successful design, a strategy based on “the central valine” concept was followed. This strategy is based on the use of a planar and unsaturated core moiety that interacts through Van der Waals interactions with a valine residue (V93) which occupies the central position of the p53 binding pocket of MDM2 (Figure 138; red arrow). Once this core is in the right position it provides with a suitable tool to successfully use the three sub-pockets present in p53.<sup>279</sup> A five-membered lactam ring, when fused to an aromatic ring is flat and able to project two exit vectors to the Phe19 pocket of MDM2 (Figure 141; magenta arrows). For this purpose, several analogues with different hydrophobic substituents in position 3 of the pyrazole ring were designed, of those the isopropyl gave the best potency.

A feature of the Leu26 sub-pocket pharmacophore was previously identified in different series of inhibitors. The  $\pi$ - $\pi$  stacking interaction between a chlorophenyl ring of the ligand and residue H96 (Figure 141; orange arrow) must be parallel for high efficiency.<sup>111, 112</sup>

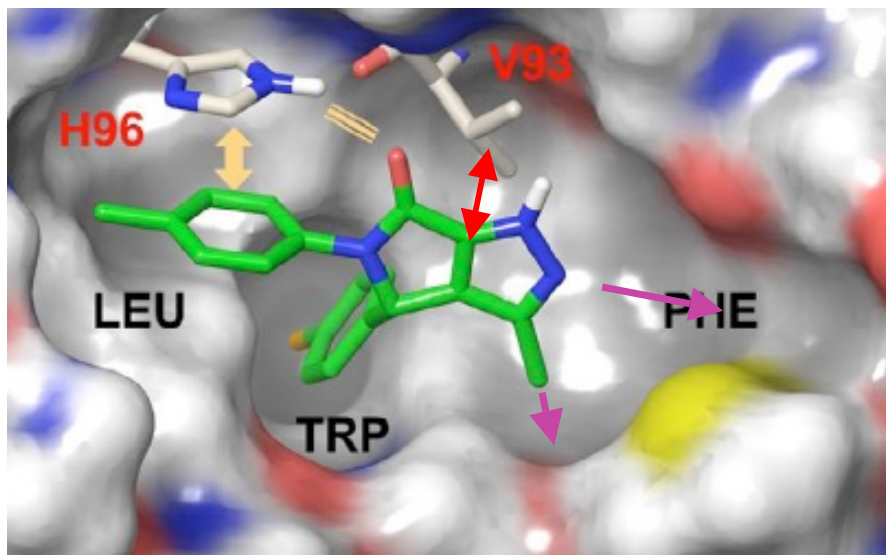


Figure 141; Crystal structure of MDM2 in complex with the pyrazolopyrrolidine (PDB ID: 5LN2)

It was speculated that the compounds could be binding with a conformational energy penalty. Modelling experiments elucidated that the introduction of two methyl groups causing steric hindrance would be adequate to avoid the conformational penalty (Figure 142) associated to the leucine and tryptophan sub-pockets.

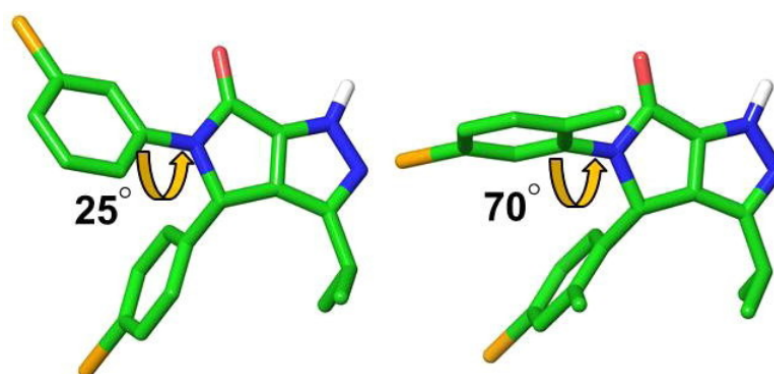


Figure 142; Calculated (*ab initio*) lowest energy conformations

After synthesizing a series of analogues exploiting the top exit vectors of the core moiety, it turned out that only one variation of this type, attachment of an ortho-methoxyphenyl group, produced a significant increase in potency. The appropriate diastereoisomer of the inhibitor (Novartis-3) showed a great final potency (Figure 143) with an IC<sub>50</sub> value of 0.13 nM in the

TR-FRET assay. An amino solubilizing chain replacing the methyl group in *ortho* position of the *meta*-chlorophenyl group was introduced to obtain a co-crystal by increasing the solubility of the compound. This modification gives us key information about the tethering position for the linker.

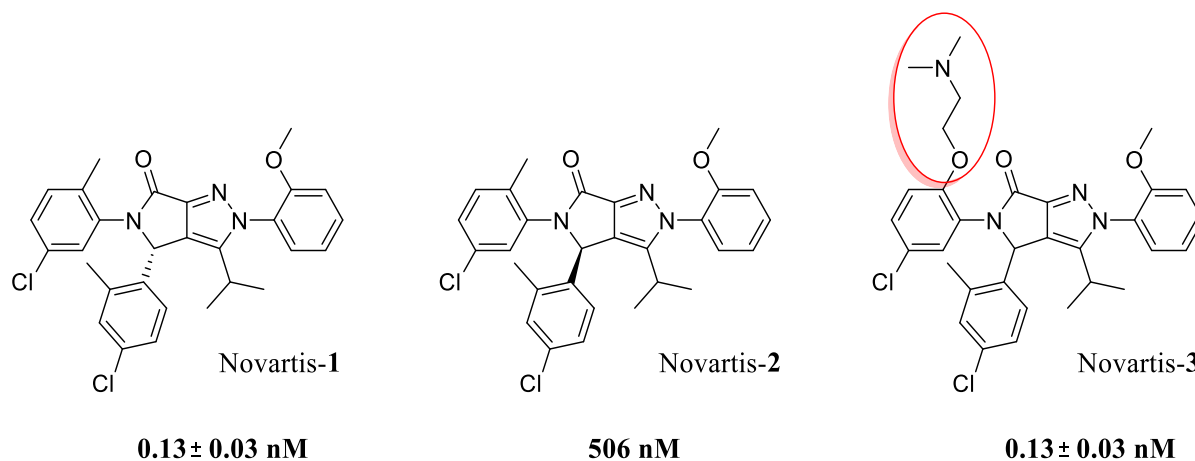


Figure 143; IC<sub>50</sub> values (nM) for the different isomers of the MDM2 inhibitor

To avoid disruption on the binding mode of the E3 ligase ligand with MDM2 (E3 ligase) the linker will be placed in a solvent exposed part of the E3 ligase ligand. The crystal structure (Figure 144) shows the dimethyl-aminoethyl substituent is exposed to the solvent, and it appears to be an ideal attachment point for the linker.

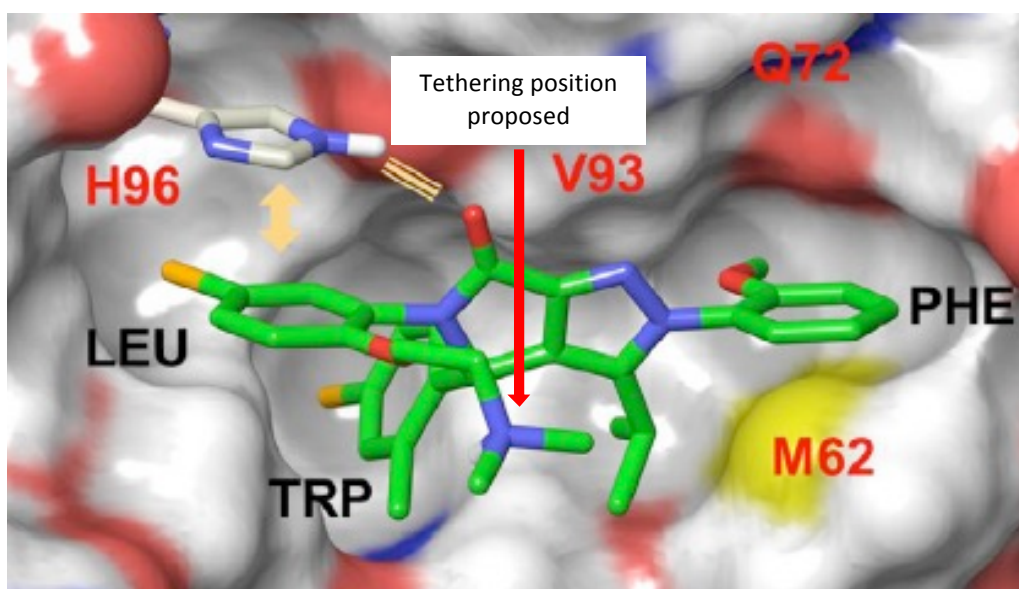


Figure 144; Crystal structure of MDM2 in complex with the inhibitor



## Chapter 21: Results and Discussion

### 21.1 Methodology and Synthesis

PROTACs are hetero-bifunctional molecules and the synthesis will be divided in three independent pathways (Figure 146).

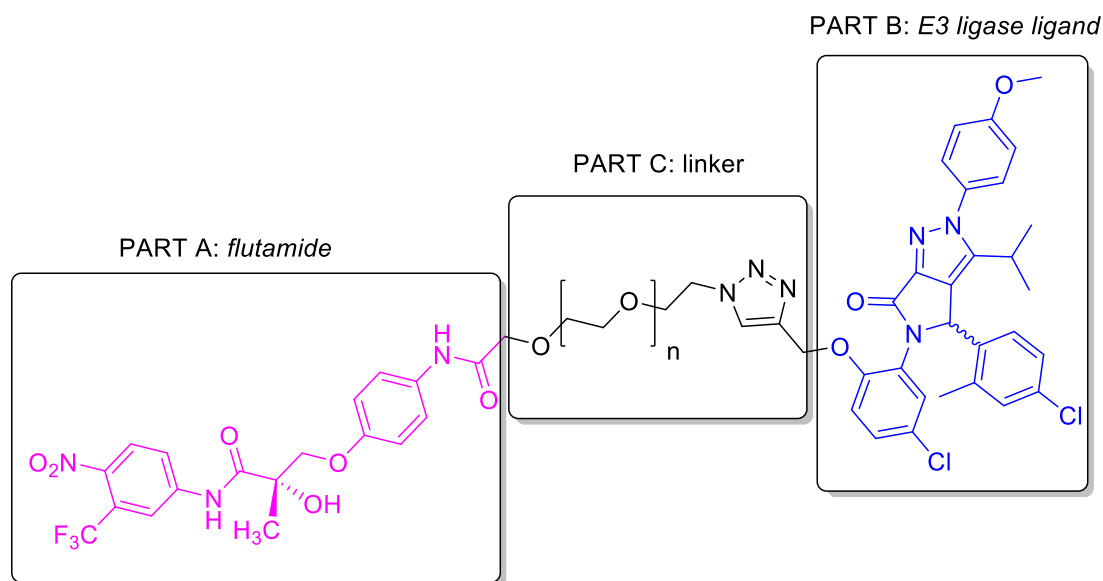
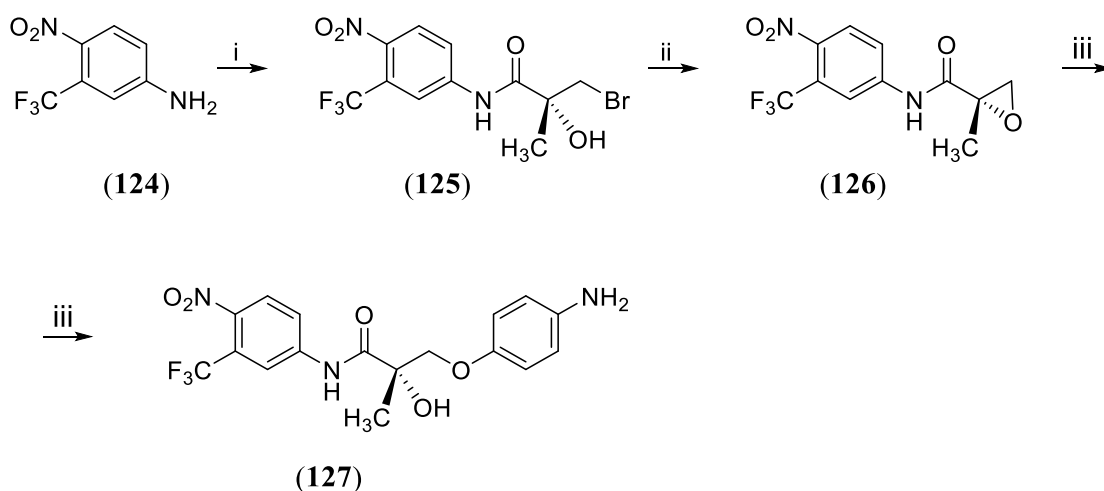


Figure 146; Molecule divided by functionalities. Three different synthetic pathways will be designed

#### 21.1.1 Part A: flutamide

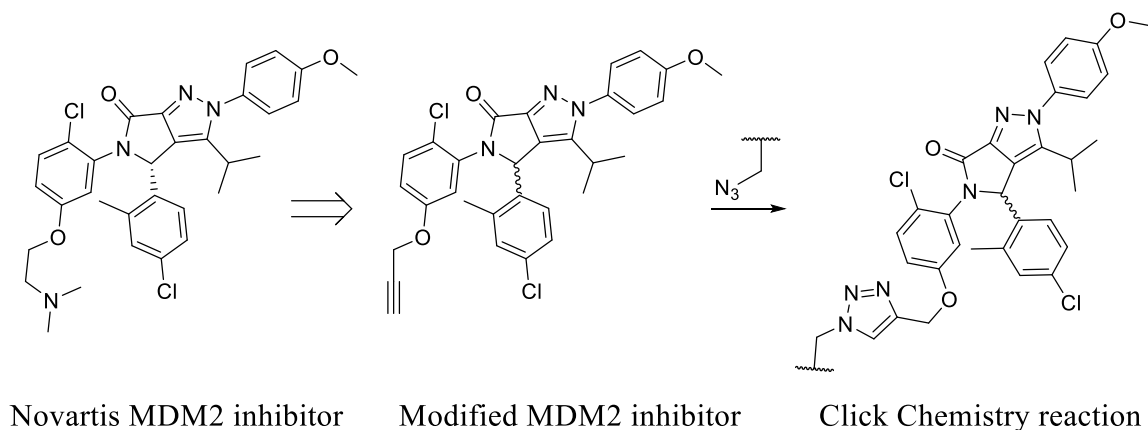
Part A of the molecule, and common to all the different targets synthesized, will be flutamide, a known inhibitor of the ARs. We will use the best analogue studied and synthesized up to date (**127**). The synthesis of Part A of the molecule starts with the amide coupling of 4-nitro-3-(trifluoromethyl)aniline (**124**) and (2*R*)-3-bromo-2-hydroxy-2-methylpropanoic acid. To do this coupling reaction an acyl chloride was prepared by using thionyl chloride and subsequent addition of the aniline (**124**) to afford compound (**125**). Then the molecule was treated with sodium hydride at 0°C in anhydrous tetrahydrofuran and then stirred at room temperature for 6 hours to yield compound (**126**). The epoxide was treated with potassium carbonate in acetone and then reacted with *N*-Boc-4-aminophenol to give us the Boc-protected desired product, which was treated with a mixture of TFA/DCM (1:1) to finally give us the desired product (**20**) (Scheme 47)



Scheme 47; i) (2*R*)-3-bromo-2-hydroxy-2-methylpropanoic acid, SOCl<sub>2</sub>, DMF, 0°C to r.t., 5h, ii) NaH(60%), THF, 0°C to r.t., 6h, iii) a) K<sub>2</sub>CO<sub>3</sub>, acetone, 65°C, 48h, b) TFA/DCM (1:1), r.t., 8h

### 21.1.2 Part B: E3 ligase ligand

Part B of the molecule is the MDM2 inhibitor synthesized by Novartis. The necessity of including a functional group that will allow us the attachment of a linker to finally build our PROTAC molecule was taken into account. In order to have a suitable and reactive end in this part of the molecule a terminal alkyne was planned for a future reaction with an azide via click chemistry. Thus, a modified version of the MDM2 inhibitor was designed (Scheme 48 – Modified MDM2 inhibitor).

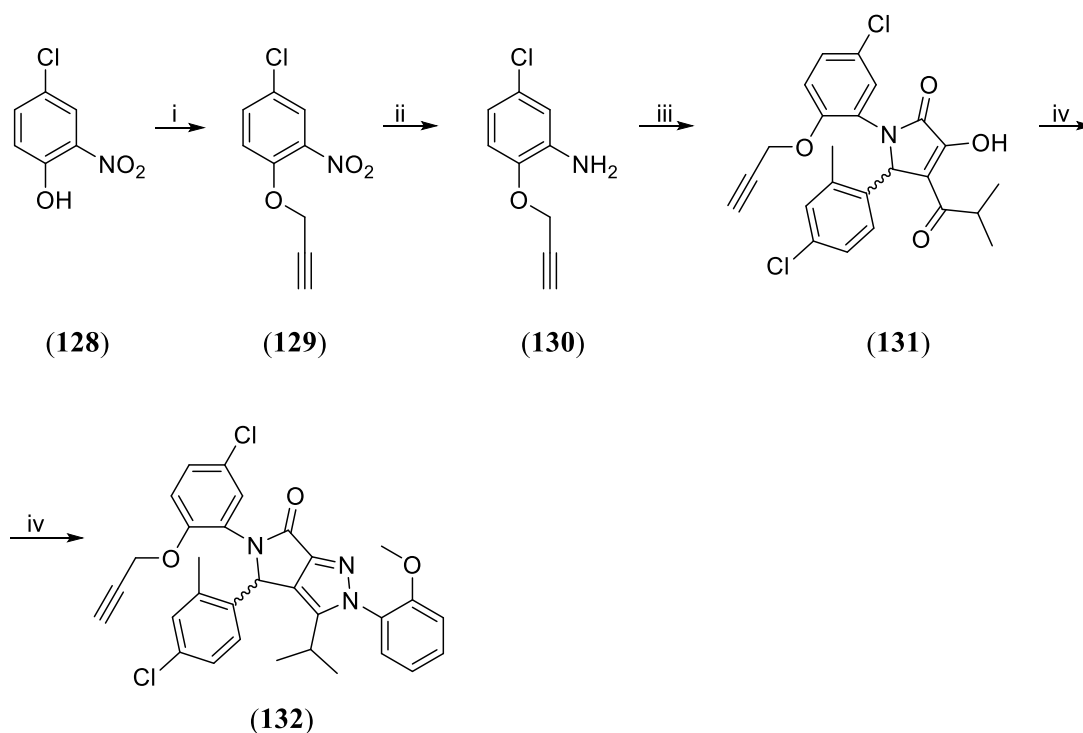


Scheme 48; Modified version of the MDM2 inhibitor for click chemistry (Cu catalysed)

The synthesis of Part B of the molecule starts with the treatment of the nitrophenol (128) with potassium carbonate and subsequent reaction with propargyl bromide to give compound (129). Then the intermediate was subjected to a reduction with SnCl<sub>2</sub>·2H<sub>2</sub>O in Ethanol under acidic conditions (conc. HCl) to yield compound (130) as a mixture with the dehalogenated



starting material, which was separated by flash column chromatography. With compound (**130**) in hand a reaction with 4-chloro-2-methyl benzaldehyde and ethyl 5-methyl-2,4-dioxohexanoate in acetic acid was performed at 80°C for 20 hours to give us the desired intermediate (**131**). Further treatment of this intermediate with 2-methoxyphenyl hydrazine hydrochloride in ethanol, using conc. HCl as a catalyst, finally gave us the modified version of the MDM2 inhibitor (**132**) (Scheme 49).



Scheme 49; i) Propargyl bromide,  $K_2CO_3$ , DMF, 0°C to r.t. 16h, ii)  $SnCl_2 \cdot 2H_2O$ , EtOH, conc. HCl, iii) 4-chloro-2-methyl benzaldehyde, ethyl 5-methyl-2,4-dioxohexanoate, acetic acid, 80°C, 20h, iv) 2-methoxyphenyl hydrazine hydrochloride, ethanol, conc. HCl (catalytic), 80°C, 48h

### 21.1.3 Part C: linker

Part C of the molecule is the linker. In this case we used PEG linkers with different lengths. The nature of the moiety to bind the linker to the different ligands was explored, using amide coupling or click chemistry reactions to yield triazoles (Figure 147).

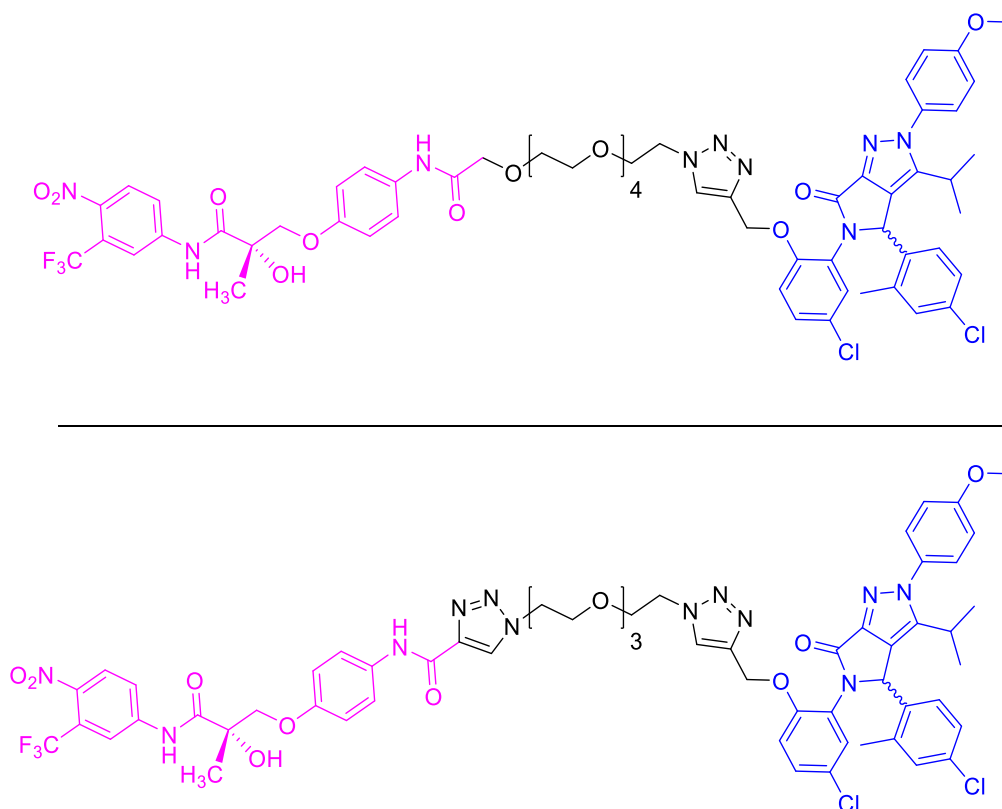
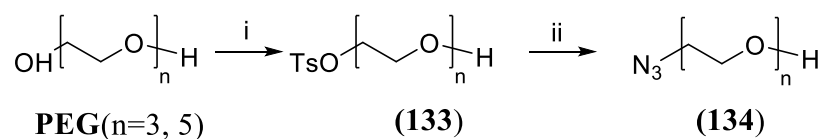


Figure 147; PROTACs to be synthesized (variations of the linker type and length))

The synthesis of part C of the molecule starts with the different polyethylene glycols and the addition of *para*-toluenesulfonyl chloride to mono-tosylate one end of the molecule to give intermediate (**133**). To avoid di-tosylation the number of equivalents of *p*-TsCl goes only up to 0.45 equivalents. After the mono-tosylation a reaction with sodium azide in DMF was performed affording the desired compound (**134**) (Scheme 50).

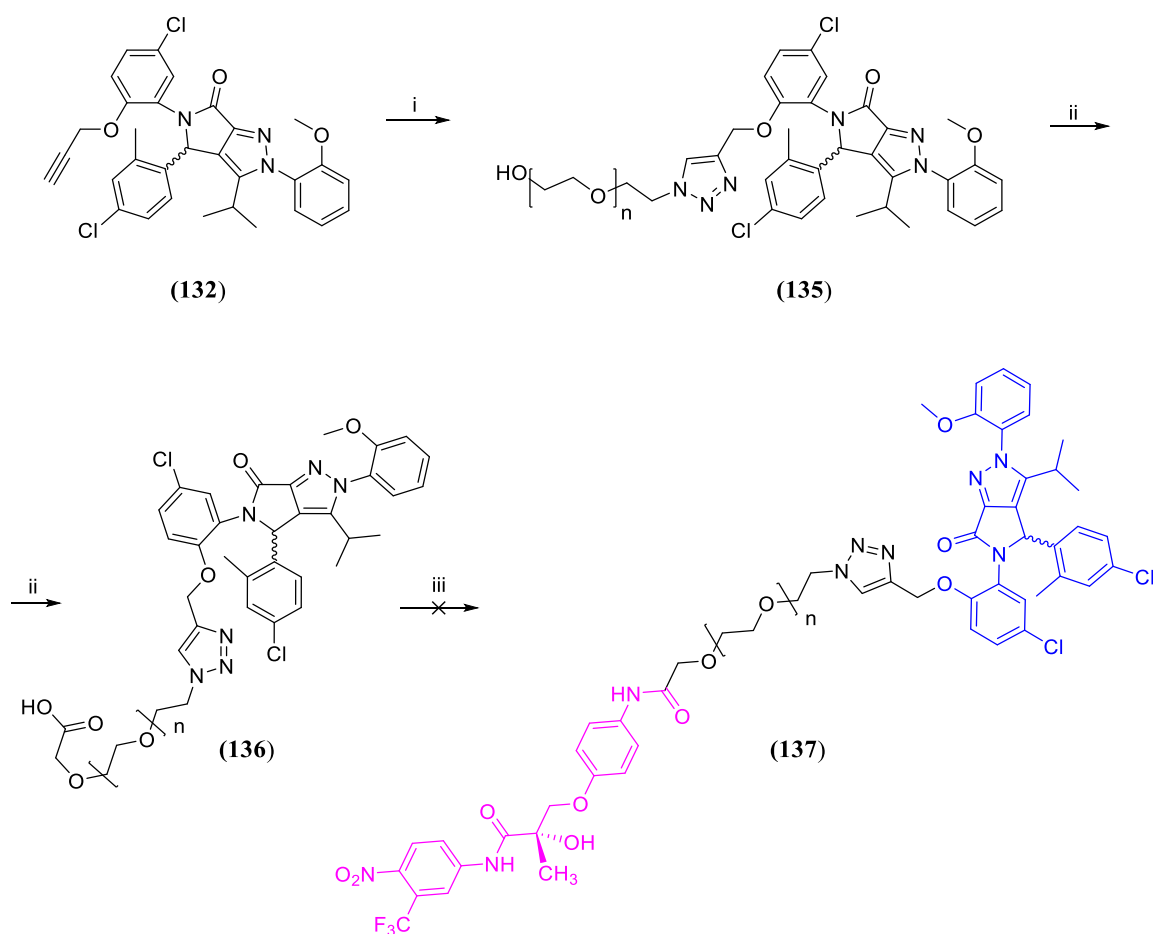


Scheme 50; i) *p*-TsCl, TEA, DCM, r.t., 16h; ii) NaN<sub>3</sub>, DMF, r.t., 16h

#### 21.1.4 Tethering of the different parts

To link all the different parts of the potential PROTAC a synthetic route was planned (Scheme 51). As a first step, the linker with a terminal azide (**134**) was reacted with the modified version of the MDM2 inhibitor (**132**) under click conditions to give (**135**). This intermediate was oxidized using Chromo Jones conditions to give (**136**). Coupling the acid with the aniline of flutamide (**127**) to give the desired PROTAC (**137**) was attempted under a

range of conditions (Table 11).



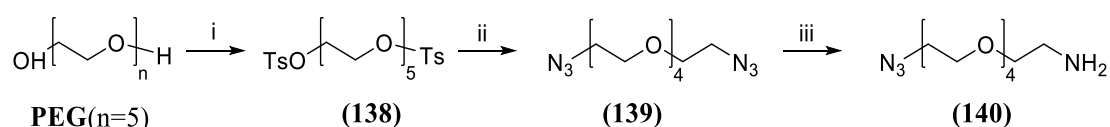
Scheme 51; i) **(134)**, (+)-Na-L-ascorbate, CuI(I), THF, 110°C, 1h, MW irradiation, ii) Chromo Jones sol. acetone, 0°C to r.t., 3h, iii) see table 5, **(127)**, r.t. to 80 °C, 16h

Compounds	Coupl. Agent	Base	Conditions	Conversion
<b>(136)</b> and <b>(127)</b>	HATU	DIPEA	Increasing Temperature and addition of crown ether	0% by LCMS
<b>(136)</b> and <b>(127)</b>	HBTU	DIPEA	Increasing Temperature and addition of crown ether	0% by LCMS
<b>(136)</b> and <b>(127)</b>	EDCI.HCl	Et <sub>3</sub> N	Increasing Temperature and addition of crown ether	0% by LCMS
<b>(136)</b> and <b>(127)</b>	SOCl <sub>2</sub>	Et <sub>3</sub> N	Increasing Temperature and addition of crown ether	0% by LCMS

Table 11; Different conditions trialled for the amide coupling

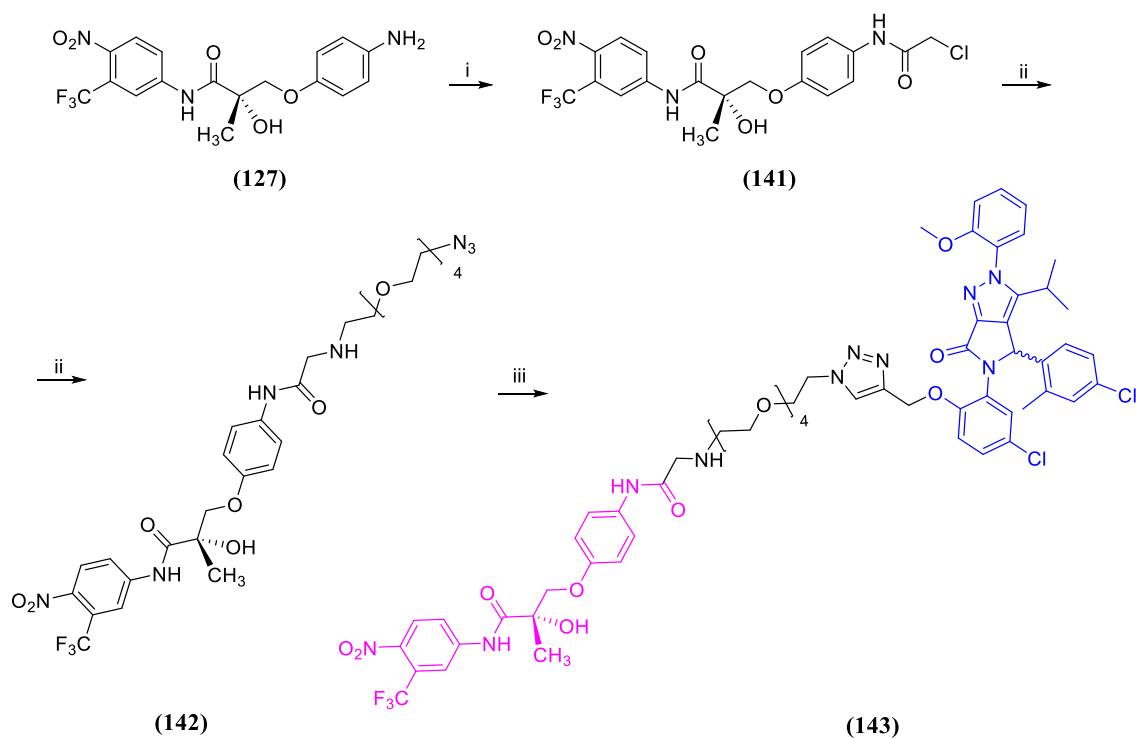
Unfortunately, the amide coupling between compound **(136)** and the aniline of compound **(127)** did not work. It was theorized that the nature of the linker could act as crown ether, coordinating with cations present in the reaction, but the addition of different crown ethers did not improve the outcome of the reaction.

A new strategy reversing the amide coupling was designed to try to synthesize the desired PROTAC molecule. For this new approach a different linker was synthesized. The tosylation of the polyethylene glycol is achieved with an excess of *p*-TsCl to give **(138)**. Then a treatment with an excess of NaN<sub>3</sub> gave **(139)**, then a Staudinger reaction gave **(140)** (Scheme 52).



Scheme 52; i) *p*-TsCl, TEA, DCM, r.t., 16h; ii) NaN<sub>3</sub>, DMF, r.t., 16h, iii) Ph<sub>3</sub>P, DCM, H<sub>2</sub>O (cat.), r.t., 5h

In this case, we did not rely on an amide coupling reaction. Instead, a  $\alpha$ -carbonyl halide SN<sub>2</sub> reaction was performed (Scheme 53- step ii) before the click reaction, which provided us with the desired PROTAC (Scheme 53 – step iii).



Scheme 53; i) Chloroacetyl chloride, K<sub>2</sub>CO<sub>3</sub>, acetone, r.t., 4h; ii) **(140)**, K<sub>2</sub>CO<sub>3</sub>, DMF, 80°C, 8h, iii) **(132)**, Na ascorbate, Cu(I)I, TBTA, 100°C, 1h, MW irradiation

Unfortunately, the presence of an epimeric mixture (Figure 148) gave us very unusual splitting patterns and duplicated signals in the  $^1\text{H}$ -NMR (Figure 149-a) and also in the  $^{13}\text{C}$ -NMR (Figure 149-b). The presence of both compounds was confirmed by HRMS.

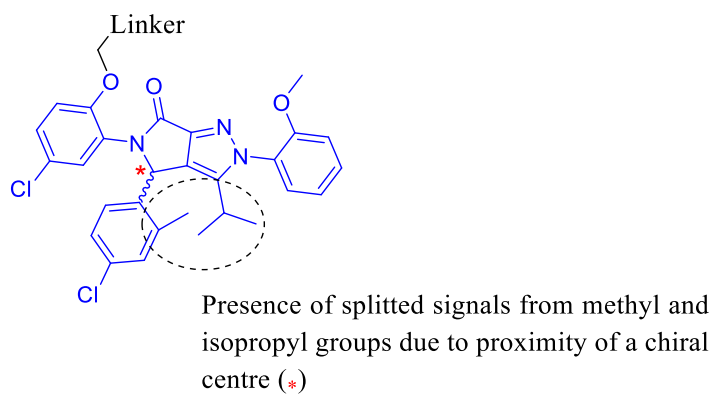


Figure 148, Circled groups with unusual splitting signals detected by NMR

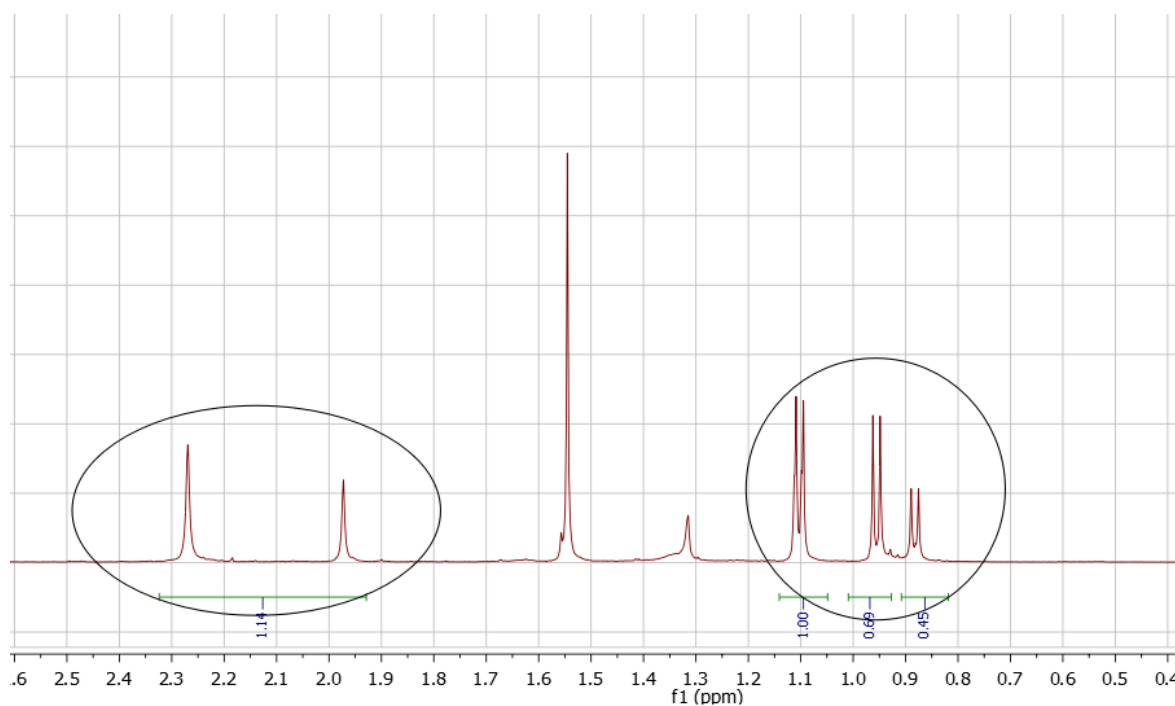


Figure 149-a;  $^1\text{H}$ -NMR, (0.8 – 1.2 ppm) for isopropyl group and (1.9 – 2.3 ppm) for aromatic methyl group

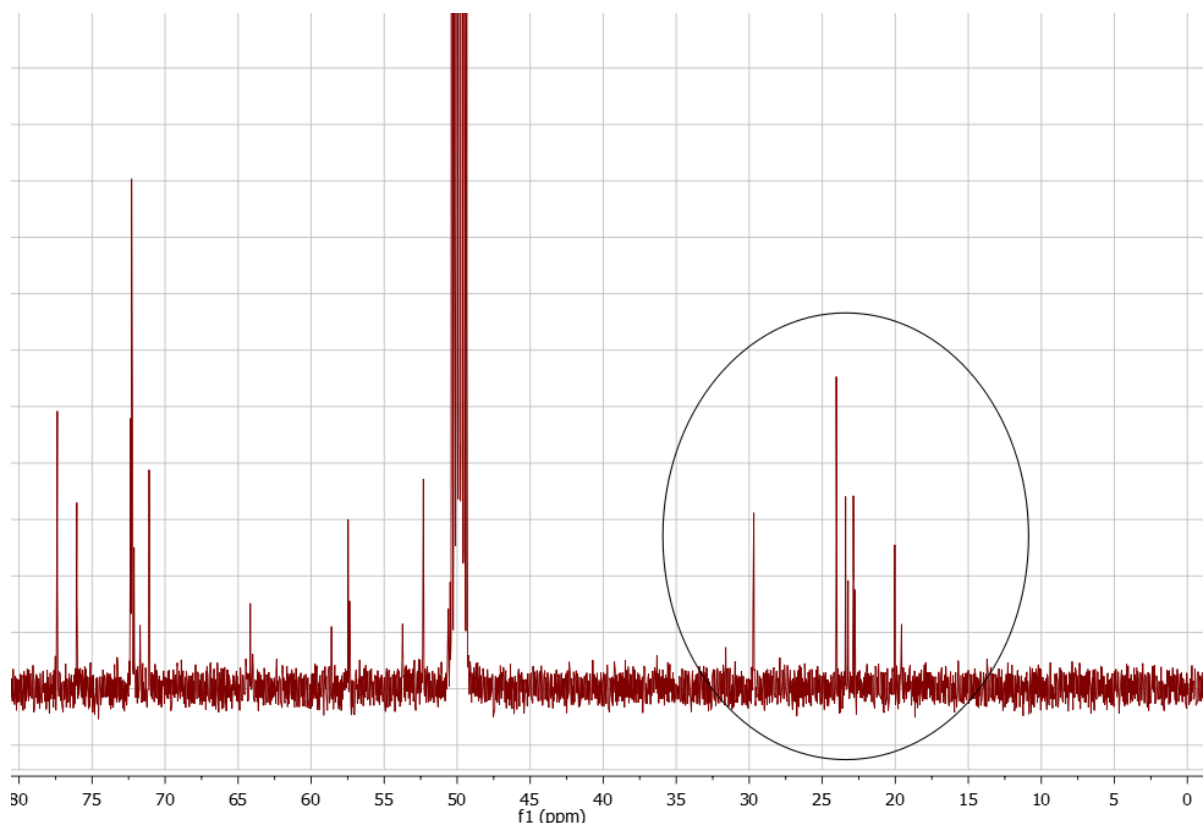
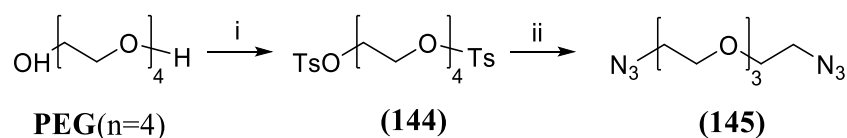


Figure 149-b;  $^{13}\text{C}$ -NMR, signals from 15 – 35 ppm for isopropyl and aromatic methyl group

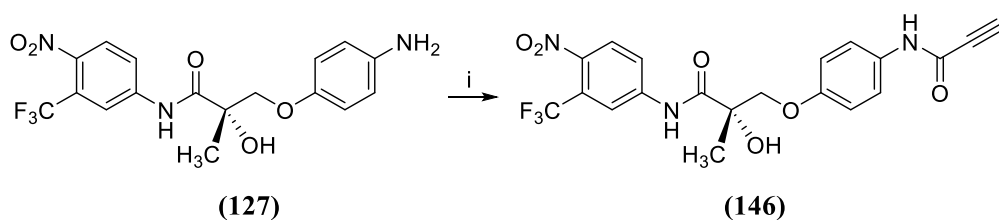
A new PROTAC with a different linker length and, at the same time, with a different moiety attaching the linker to the flutamide and to the E3 ligase was also designed. For this, we used a di-azide as a linker and two click reactions were performed to join the two ligands and the linker. In one end with a new version of flutamide, functionalized with an alkyne (**146**) and on the other end the MDM2 inhibitor (**132**).

For the synthesis of the new linker a di-tosylation followed by treatment with excess of sodium azide was performed to yield (**145**) (Scheme 54).



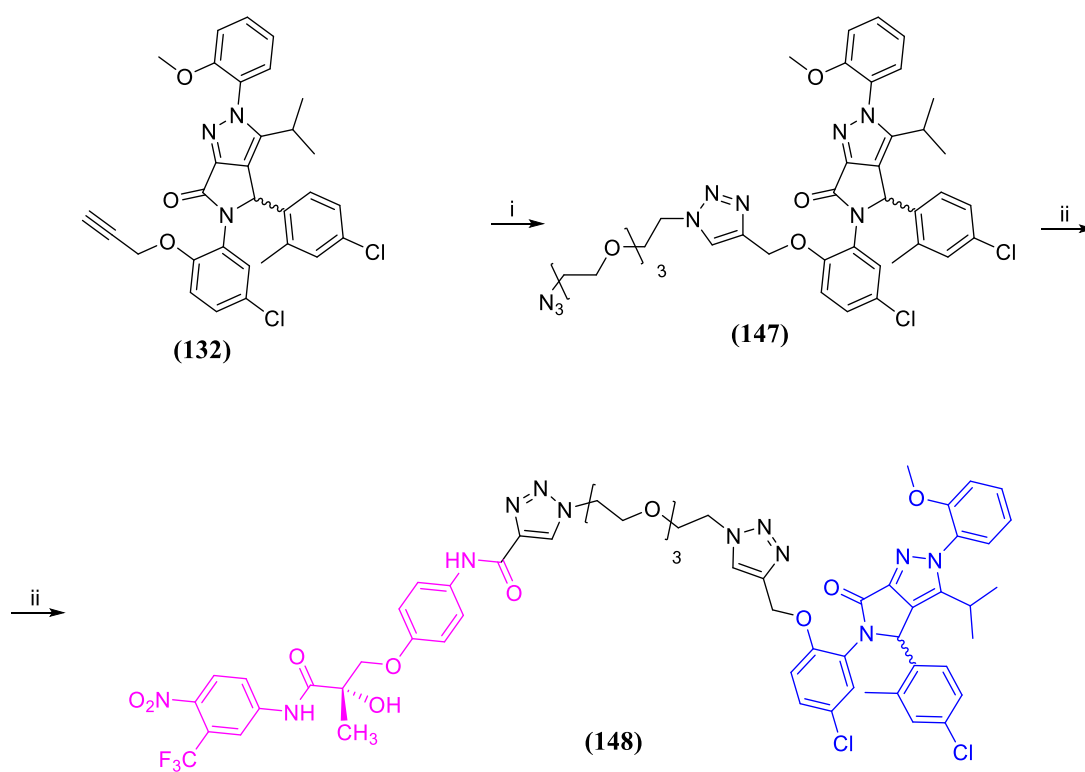
Scheme 54; i) *p*-TsCl, TEA, DCM, r.t., 16h; ii) NaN<sub>3</sub>, DMF, r.t., 16h

For the synthesis of the new version of flutamide, compound (**127**) was used to perform an amide coupling reaction using propiolic acid (Scheme 55).



Scheme 55; i) Propiolic acid, EDCl.HCl, TEA, THF, r.t., 16h

The click chemistry was performed first with compound (146) and then with compound (145) using similar conditions in both cases (Scheme 58). To try to entropically minimize the formation of homo-bivalent compounds in the first step (Scheme 56-step i) an excess of linker was used.



Scheme 56; i) (139), Na ascorbate, Cu(I)I, TBTA, 100°C, 1h, MW irradiation; ii) (146), Na ascorbate, Cu(I)I, TBTA, 100°C, 1h, MW irradiation

Again, the presence of an epimeric mixture (see Figure 148) gave us very unusual splitting patterns and duplicated signals in the  $^1\text{H}$ -NMR (Figure 150-a) and also in the  $^{13}\text{C}$ -NMR (Figure 150-b). The presence of both compounds was confirmed by HRMS.

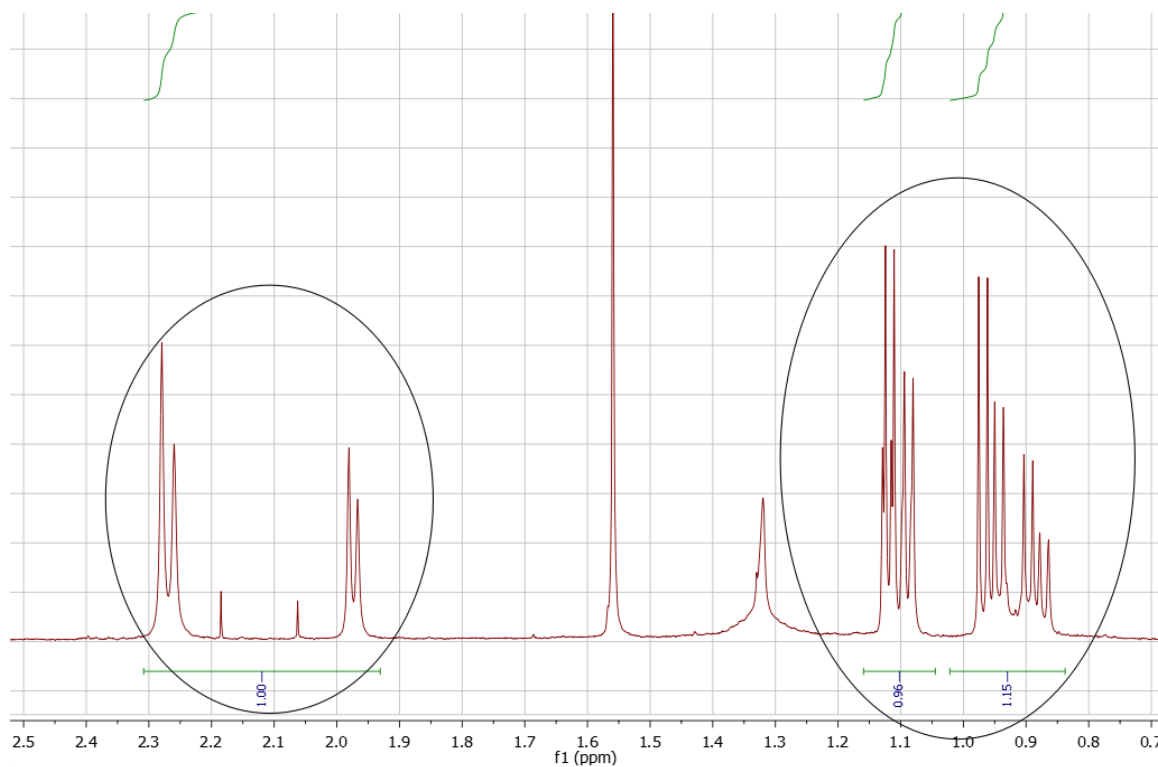


Figure 150-a;  $^1\text{H}$ -NMR, (0.8 – 1.2 ppm) for isopropyl group and (1.9 – 2.3 ppm) for aromatic methyl group

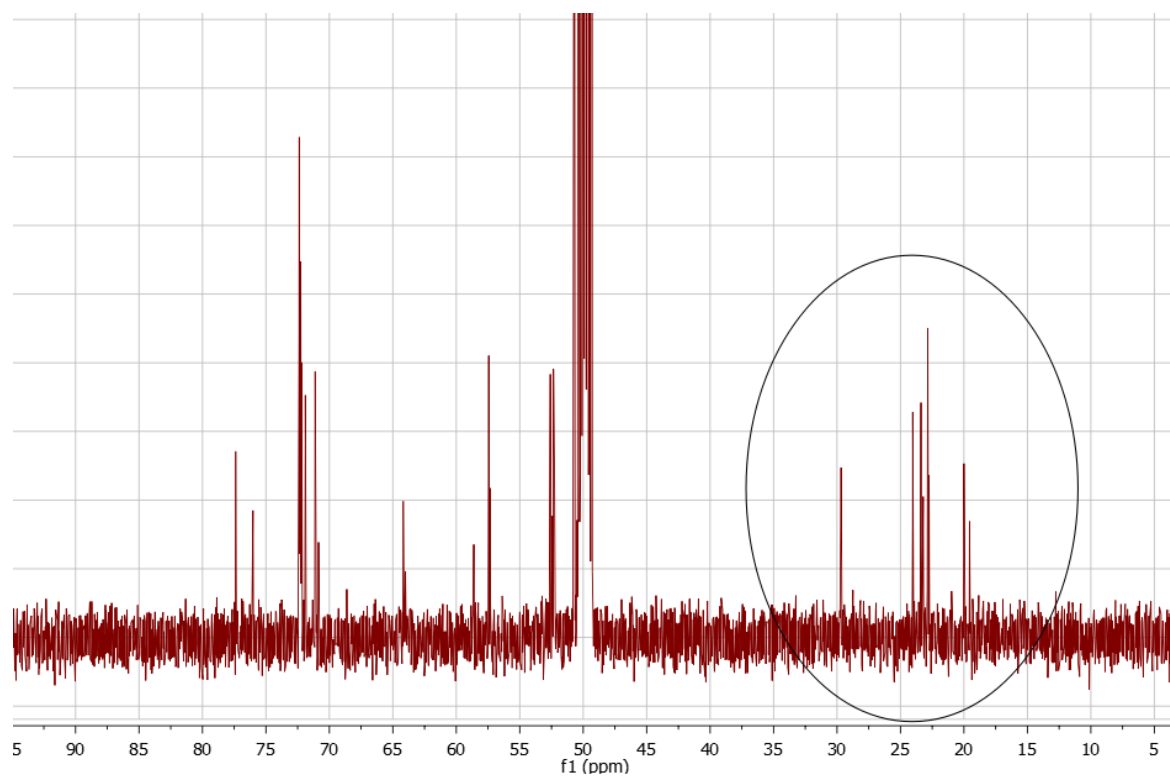
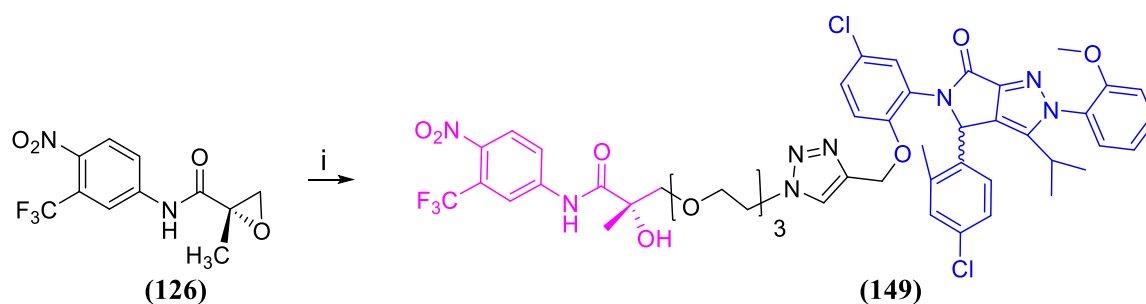


Figure 150-b;  $^{13}\text{C}$ -NMR, signals from 15 – 35 ppm for isopropyl and aromatic methyl group



As the linker chemistry proved to be challenging an additional PROTAC was also developed with a less potent version of flutamide. In this case the primary alcohol of an ethylene glycol linker was used as a nucleophile to open the cyclopropoxy ring present in compound **(126)** (Scheme 57) with the use of sodium hydride to deprotonate the terminal alcohol of the linker.



Scheme 57; i) **(135)** ( $n = 2$ ), NaH (60%), THF, 0°C to r.t., 16 h

The presence of an epimeric mixture (see figure 148) should have given us unusual splitting patterns and duplicated signals in the  $^1\text{H}$ -NMR (Figure 151-a) and  $^{13}\text{C}$ -NMR (Figure 151-b), but in this case the spectrum was well defined.

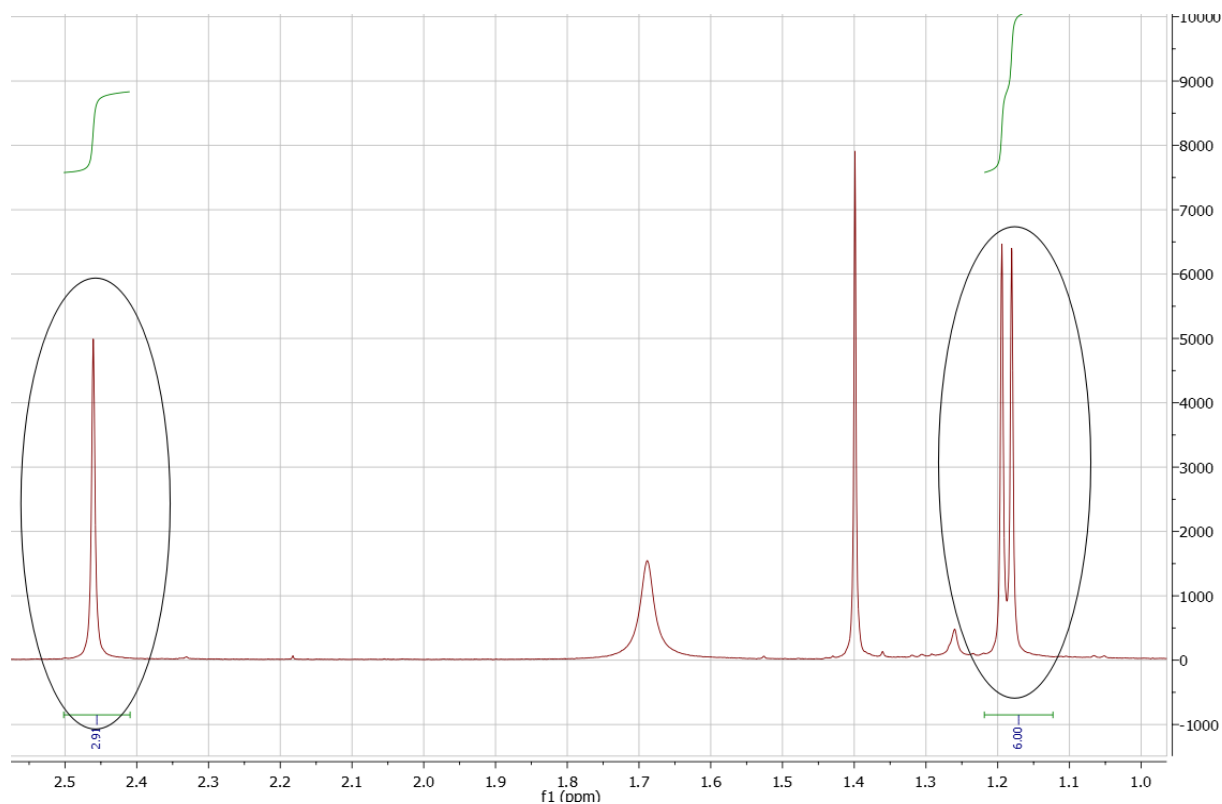


Figure 151-a;  $^1\text{H}$ -NMR, (1.1 – 1.25 ppm) for isopropyl group and (2.4 – 2.5 ppm) for aryl methyl group

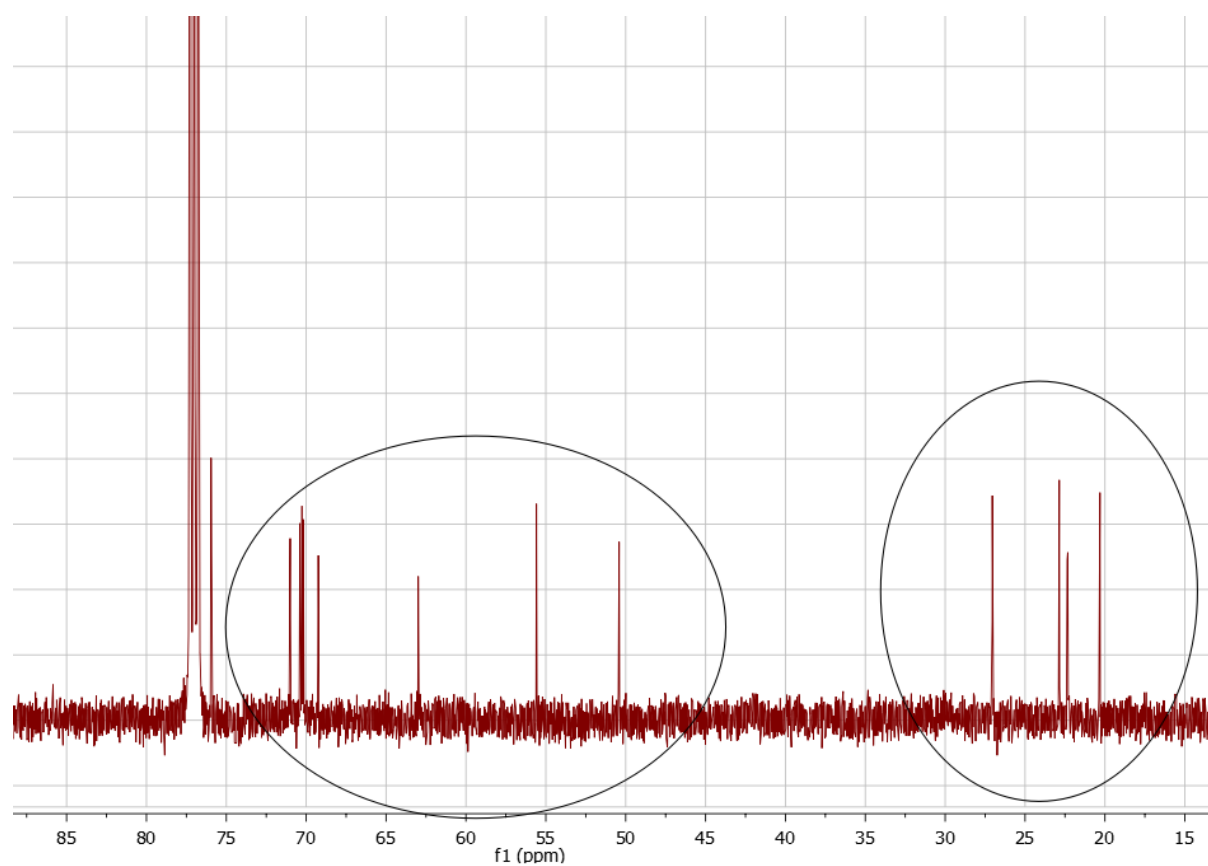


Figure 151-b;  $^{13}\text{C}$ -NMR, no duplicated signals

## Chapter 22: Biological Assays

The PROTACs were submitted for testing. The results are summarized in this chapter. The compounds selected to be tested with different AR cell lines are classified below:

Compound	Label	Structure (AR inhibitor-E3 ligase ligand)
NCL-00023900	PROTAC 1	Flutamide-isoindolinone (MDM2 ligand)
(149)	PROTAC 2	Flutamide-pyrazolopyrrolidine (MDM2 ligand)
(143)	PROTAC 3	Flutamide-pyrazolopyrrolidine (MDM2 ligand)
(148)	PROTAC 4	Flutamide-pyrazolopyrrolidine (MDM2 ligand)
547-244-1	PROTAC 5	Apalutamide derivative-pomalidomide (cereblon ligand)

Table 12; PROTACs to be tested with different AR cell lines

For structure of PROTAC 1 see Figure 149-b, for structure of PROTAC 2 (compound (149)) for structure of PROTAC 3 (compound (143)) and for PROTAC 4 (compound (148)). PROTAC 5 is not related to the previous compounds but also targets the androgen receptor and is based on Cereblon as an E3 ligase instead of MDM2 so we can compare a new E3 ligase ligand targeting the same protein, see Figure 152-a.

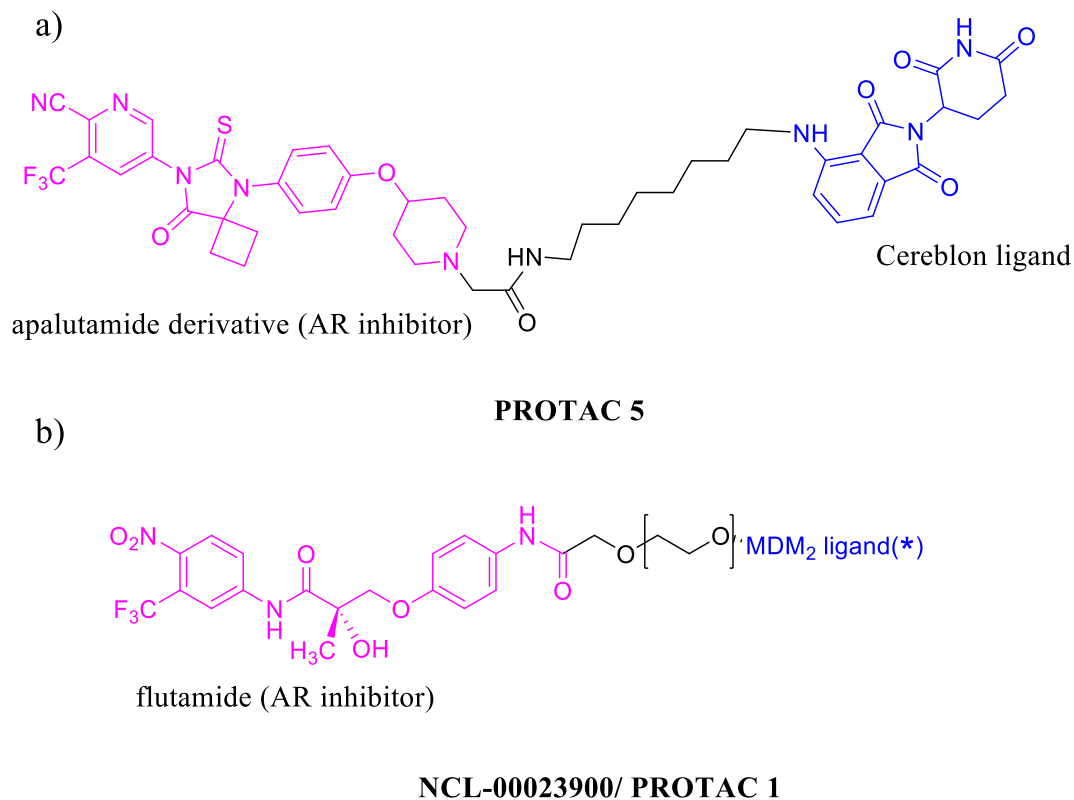


Figure 152; a) Cereblon-based PROTAC 5; b) MDM2-based PROTAC 1 (\*confidential structure)

The aim is to test the effects of the PROTACs on a panel of PCa cell lines with defined androgen receptors status upon addition of DHT and each of the PROTACs by Western blot analysis. In this case, the plating was done using several types of cell lines summarized in the table below:

<u>Cell line</u>	<u>AR mutation</u>	<u>Androgen Responsive</u>	<u>Prostate Specific Antigen (PSA)</u>
<b>LNCaP</b>	Mutation in AR ligand binding domain	Positive	Expresses PSA mRNA/protein
<b>VCap</b>	Express wild type AR and low levels of AR- V7	Weakly	Expresses PSA
<b>CWR22Rv1</b>	Express mutant (H874Y) AR and AR variants	Weakly	Secretes low levels of PSA
<b>AR-EK</b>	CRISPR modified CWR22Rv1 to remove full length AR. Expresses AR variants	Negative	Secretes low levels of PSA

Table 13; Different AR cell lines and their properties

During the analysis the levels of AR, PSA and FKBP-5 (binding protein for the role of regulation and tumourigenesis) were monitored.  $\alpha$ -tubulin was used as positive control for the analysis.

Also, one of the ways to confirm that the activity of our PROTACs is based in degradation rather than straightforward inhibition could be the addition of a proteasome inhibitor but also the confirmation of the “hook effect”. The fundament of this effect is that the proportion of target protein degraded would be reduced at high concentrations of PROTAC. Due to binding of the target by the PROTAC as a binary complex and, at the same time, binding of the E3 ligase by the PROTAC also as a binary complex, the formation of the ternary complex is not possible. Therefore, high concentrations of PROTAC block induced degradation effectively (Figure 153).

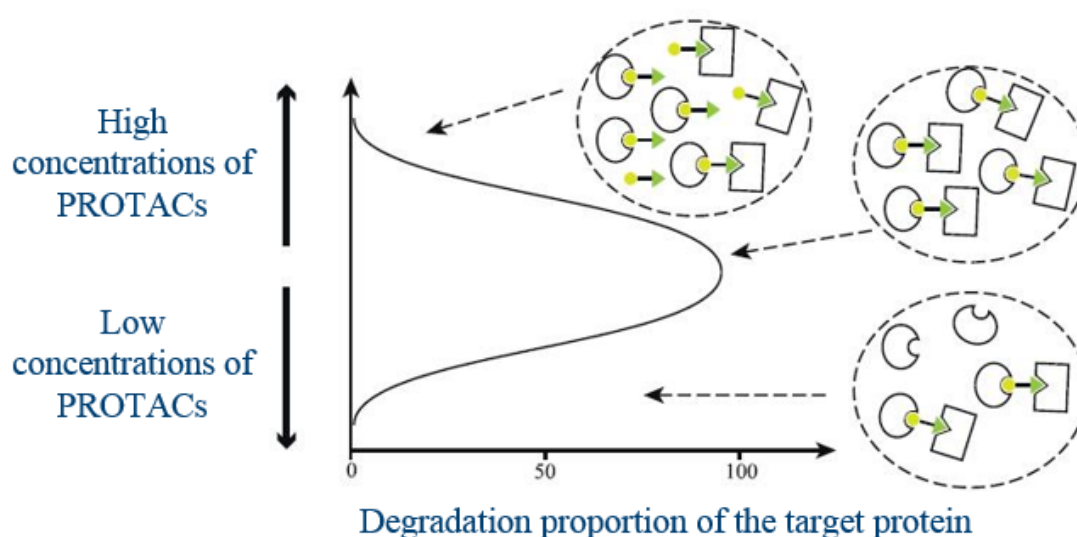


Figure 153; Hook effect (Sainan, An *et al.*, 2018)

## 22.1 Western blot analysis (enzalutimade as control)

The results and conclusions obtained after Western Blot Analysis are summarized below:

### 22.1.1 Addition of PROTACs LNCap cells

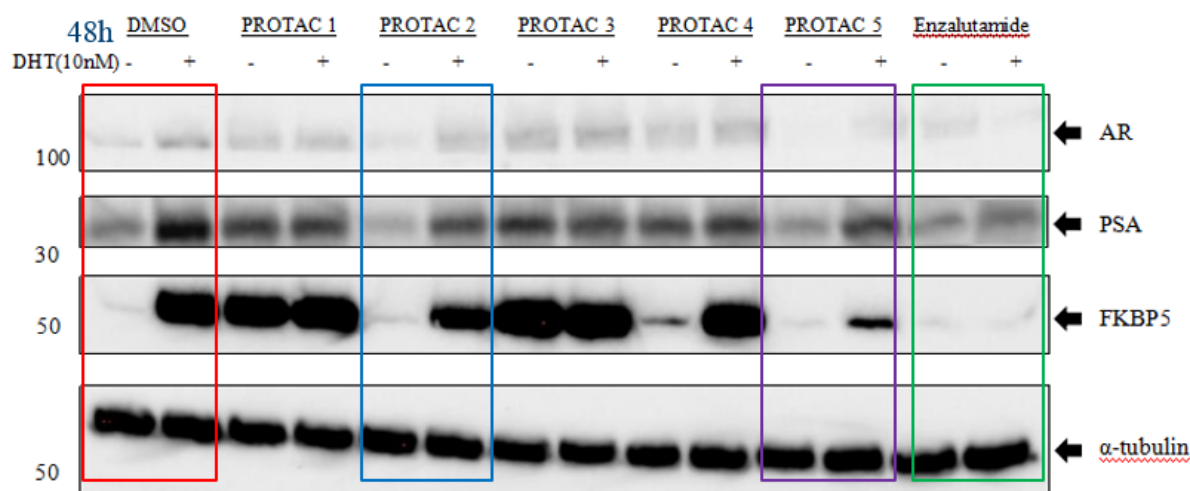


Figure 154; Western blot analysis of LNCap cells upon addition PROTACs

- In the case of the AR we can see a subtle reduction in PROTACs 2 (blue box) and 5 (purple box). For PROTAC 5, the reduction is similar to that observed for enzalutamide (green box). For PROTACs 1, 3 and 4 the level of AR shows a subtle increase in the levels of AR without DHT suggesting a possible weak agonistic activity. Unfortunately, the results observed are inconclusive. The differences in signals are extremely weak and the assumptions made here are not definite. Further

testing should be implemented specially in the case of AR for consistent interpretation.

- The observation of PSA expression levels gives us similar information, but in this case the intensity of the signals is slightly more reassuring. As we can see the PSA expression levels for PROTACs 1, 3 and 4 is enhanced in the absence of DHT addition, suggesting again a possible agonistic activity. For PROTACs 2 (blue box) and 5 (purple box) the reduction in the PSA signal upon addition of DHT is very subtle but consistent with the expression levels of androgen receptors, suggesting possible degradation or simple inhibition.
- FKBP5 levels are clearly reduced for PROTAC 5 (purple box), and subtly reduced for PROTAC 2 (blue box). Again, PROTACs 1, 3 and 4 suggest some weak agonistic activity due to the results observed in the absence of DHT.
- As a general conclusion, because LNCap Cells are androgen responsive with mutated androgen receptors, enzalutamide should not be as effective as an antagonist due to the mutation of the androgen receptor. In the case of flutamide,<sup>281</sup> agonistic activity is observed for mutated androgen receptors, and the fact that some agonistic activity is observed reflects that our PROTAC is based on flutamide, these results would suggest that flutamide engages the mutated androgen receptor with agonistic activity and that the PROTAC is not producing degradation.

### 22.1.2 Addition of PROTACs to VCap cells:

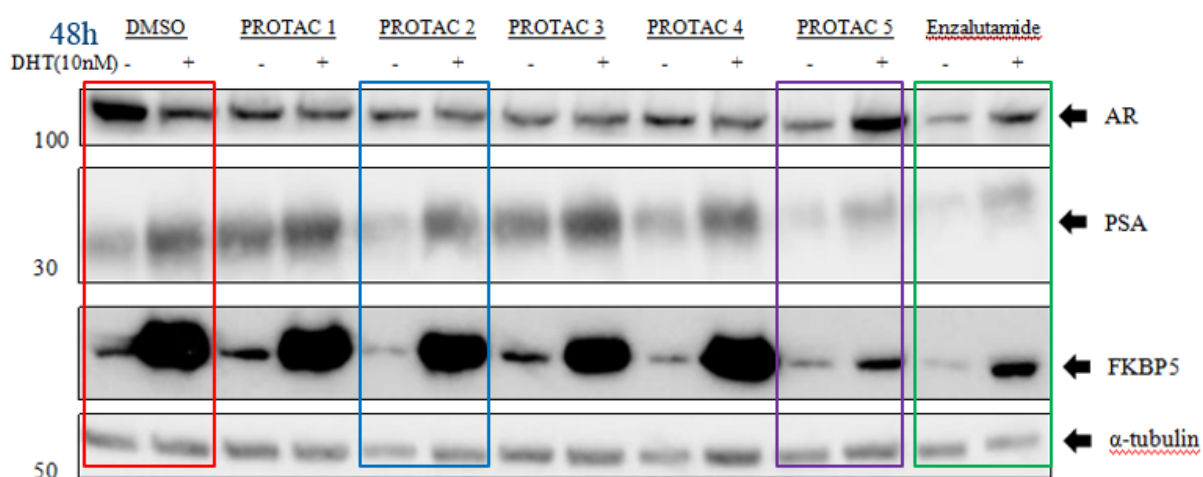


Figure 155; Western blot analysis of VCap cells upon addition PROTACs

- The results of the Androgen Receptor are inconclusive.

- In the case of the PSA expression levels we can see a subtle decrease after addition of PROTACs 2 (blue box) and 5 (purple box). PSA expression levels are similarly reduced after addition of enzalutamide. For PROTACs 1, 3 and 4 again a weak agonistic activity can be suggested from the results in the absence of DHT.
- Results obtained for FKBP5 showed a great reduction on the expression levels for the protein upon addition of DHT for PROTAC 5, similar to the result observed for enzalutamide.
- As a general conclusion, we know that androgens directly regulate FKBP5 via an interaction between the androgen receptor and a distal enhancer located downstream of the transcription start site in the fifth intron of the FKBP5 gene.<sup>282</sup> Unfortunately, this result by itself cannot confirm degradation of the androgen receptors, specially in this case, where no change is observed in the expression levels of androgen receptor. A weak agonistic activity can be inferred again for PROTACs 1 and 3, ruling out degradation, but the results are not conclusive.

### 22.1.3 Addition of PROTACs to CWR22Rv1 cells:

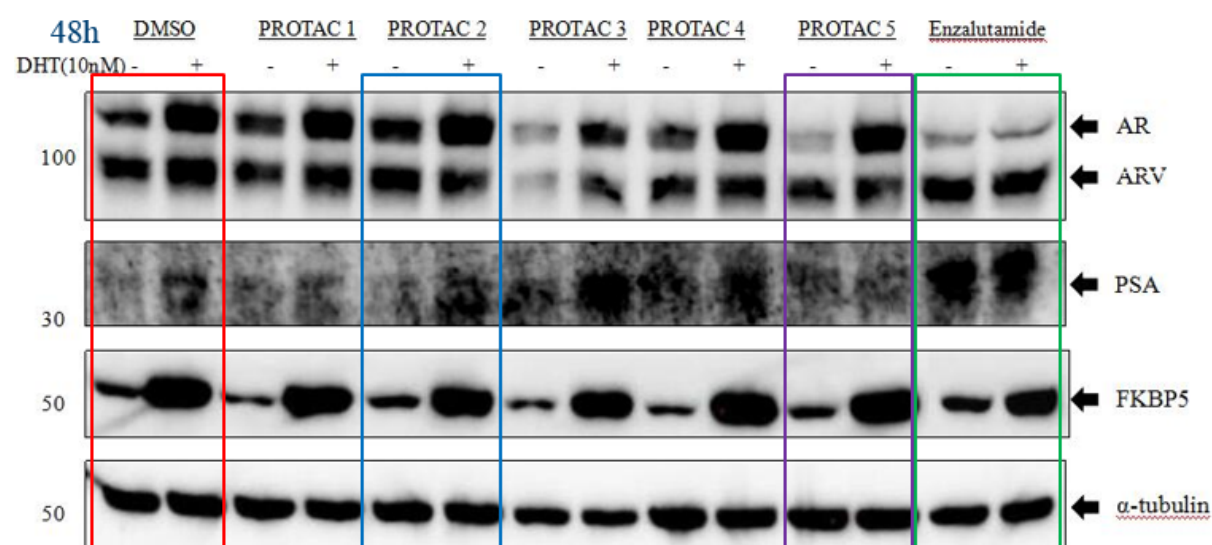


Figure 156; Western blot analysis of CWR22Rv1 cells upon addition PROTACs

- In the case of the androgen receptors and its variant, the results are not conclusive. As we can see, almost the same level of AR and ARV expression is observed, with or without DHT addition in the control experiment. In the case of PROTAC 3 a weak agonistic activity could be suggested.
- PSA cannot be interpreted.

- FKBP5 expression levels remained almost constant during our experiments, no conclusion can be obtained with these results.
- Inconclusive, no interpretation can be made with these results in hand.

#### 22.1.4 Addition of PROTACs to AR-Vs in AR-EK cells:



Figure 157; Western blot analysis of AR-EK cells upon addition PROTACs

- PROTACs (1 to 5) or flutamide are unable to target the proteins. Levels are constant during the assays. Further repetition of the assay should be implemented.
- Inconclusive results.

#### 22.2 Western blot analysis (flutamide as control)

In order to compare flutamide based PROTACs with flutamide itself a series of experiments were performed with the androgen receptor inhibitor as a control. For the immunoblotting assays see below:



### 22.2.1 Addition of PROTACs LNCap cells:

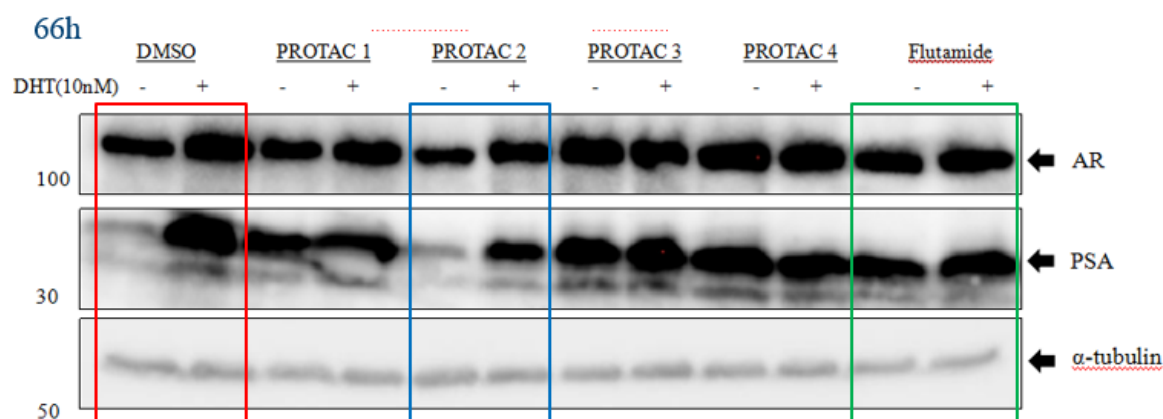


Figure 158; Western blot analysis of LNCap cells upon addition PROTACs

- In the case of PROTAC 2 we can see a reduction in the level PSA in the absence of DHT, suggesting that is the only PROTAC without an agonistic activity as we can see in the rest of them.

### 22.2.2 Addition of PROTACs to VCap cells:

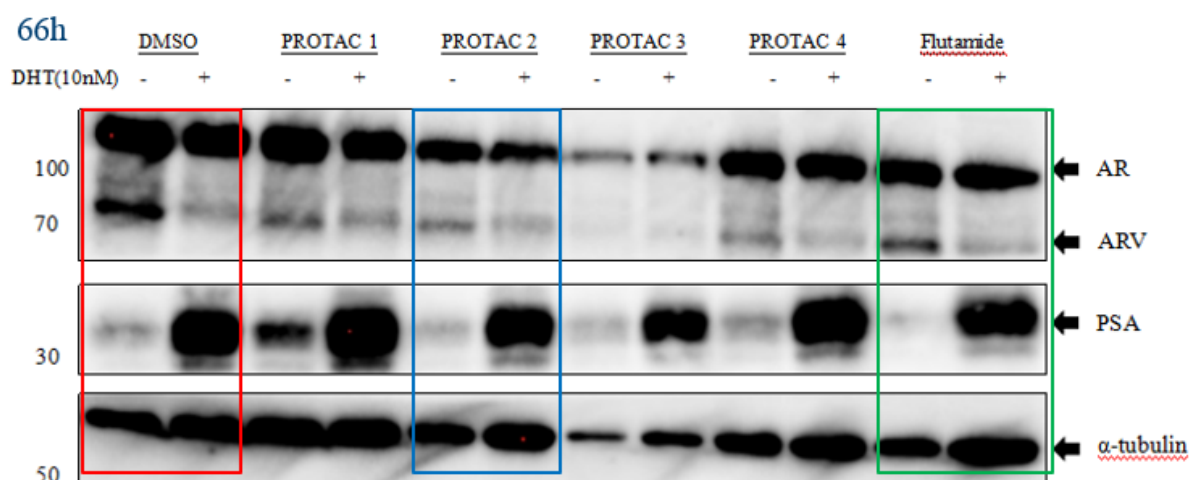


Figure 159; Western blot analysis of VCap cells upon addition PROTACs

- Results for AR and ARV are inconclusive.
- Very subtle reduction of PSA levels for PROTAC 2 (blue box).
- Antagonistic activity could be assumed for PROTACs 2 (blue box) and 3 with the reduction of PSA expression levels observed but the further experimentation would be required to reassure degradation instead of inhibition by flutamide itself.

## 22.3 Testing the hook effect

PROTACs 2 and 5 will be re-tested for the hook effect as the most promising compounds from the previous tests, engaging the androgen receptors without an agonistic effect (enzalutamide as control) and showing weak agonistic activity (flutamide as control).

### 22.3.1 Addition of enzalutamide in LNCaP and VCaP cells

Enzalutamide was used as control for the hook effect.

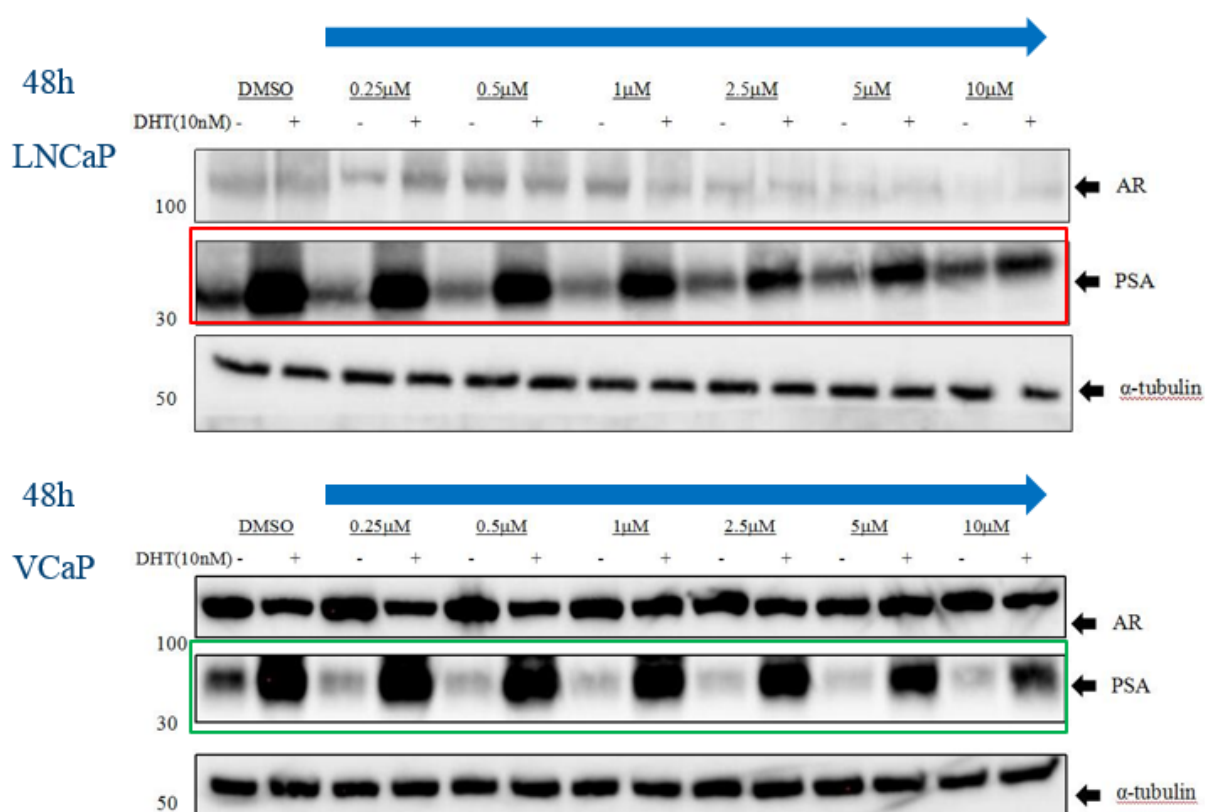


Figure 160; Western blot analysis of LNCaP and VCaP cells upon addition of enzalutamide (0.25 μM to 10 μM)

- The hook effect is not observed upon addition of enzalutamide.

As the most promising compounds from previous tests PROTAC 2 and 5 were tested in concentrations from 0 μM to 10 μM. The results are shown below:

### 22.3.2 PROTAC 5 in LNCaP and VCaP cells:

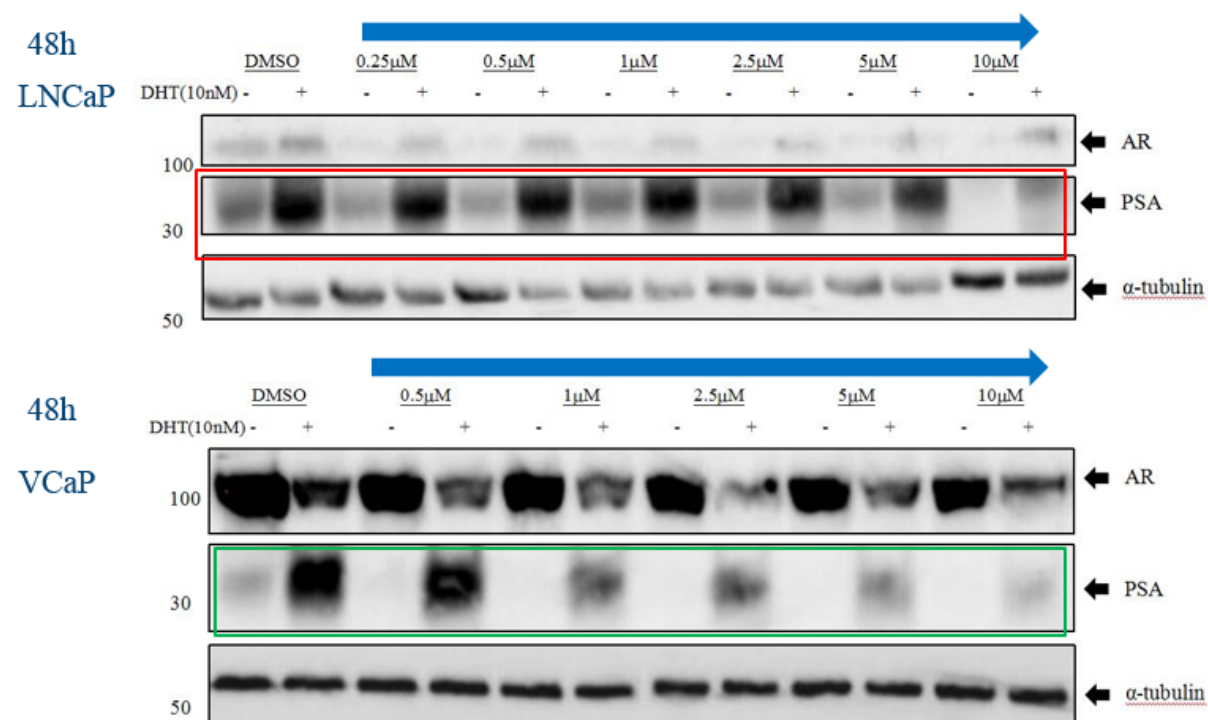


Figure 160; Dose-effect of PROTAC 5 in LNCaP and VCaP cells

- In the case of LNCaP cells a slight increase in the expression of androgen receptors levels is observed, suggesting the presence of the hook effect.
- With the increase in PROTAC 5 we can see a decrease in the PSA levels for both cell lines but no presence of the Hook Effect is observed.
- Results with androgen receptors expression levels and PSA expression levels are not consistent. Results are not conclusive.

### 22.3.3 PROTAC 2 in LNCaP and VCaP cells:

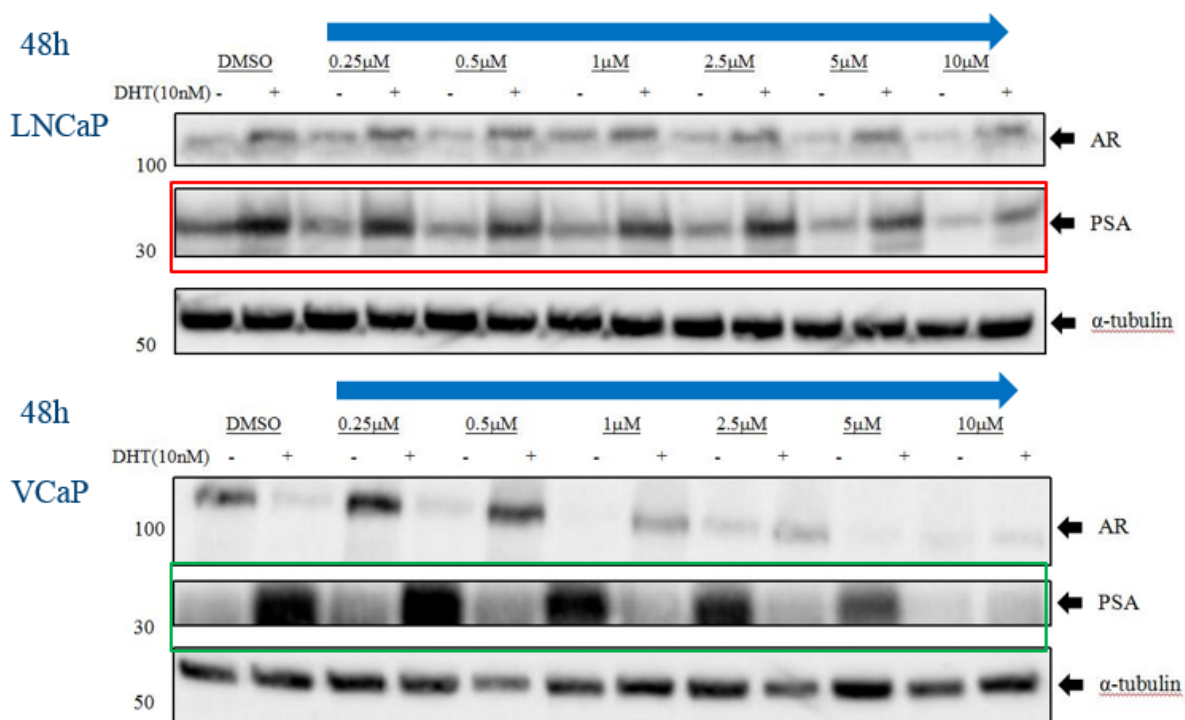


Figure 161; Dose-effect of PROTAC 5 in LNCaP and VCaP cells

- Very subtle decrease of the level of PSA in LNCaP cells is observed. No presence of hook effect.
- In the case of the VCaP cells a reduction of AR and PSA levels is observed but no presence of Hook Effect can be confirmed.

Finally, a combination effect of PROTAC 2 with nutlin-3a was proposed. According to the research performed John Lunec,<sup>283</sup> 5 μM of nutlin-3a could inhibit LNCaP proliferation by 50% with subsequent accumulation of MDM2 in the cell as a result of a positive feedback loop induced by the high levels of p53 in the cell which activates the expression of MDM2 by interaction with the *MDM2* gene.

#### 22.3.4 Addition of PROTAC 2 and Nutlin-3a to LNCaP cells:

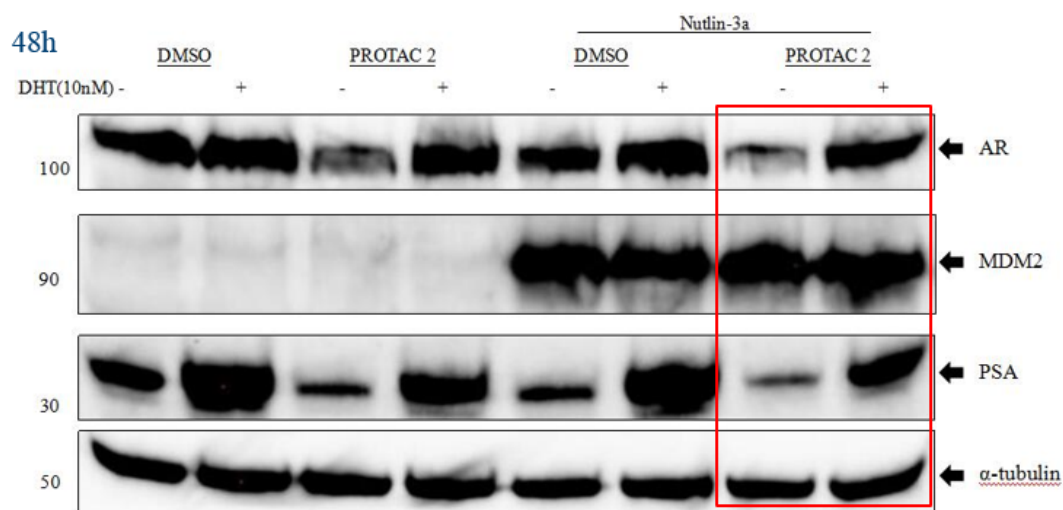


Figure 162; Addition of PROTAC 2 and Nutlin-3a to LNCaP cell

- In the case of the combination between Nutlin-3a and PROTAC 2 no further degradation of PSA or androgen receptors is observed in the combination experiment. The expression levels of androgen receptors in the control experiment are similar upon addition of DHT, suggesting that the experiment was not conducted properly. Results are inconclusive.

#### 22.3.5 Addition of PROTAC 2 and Nutlin-3a to VCaP cells

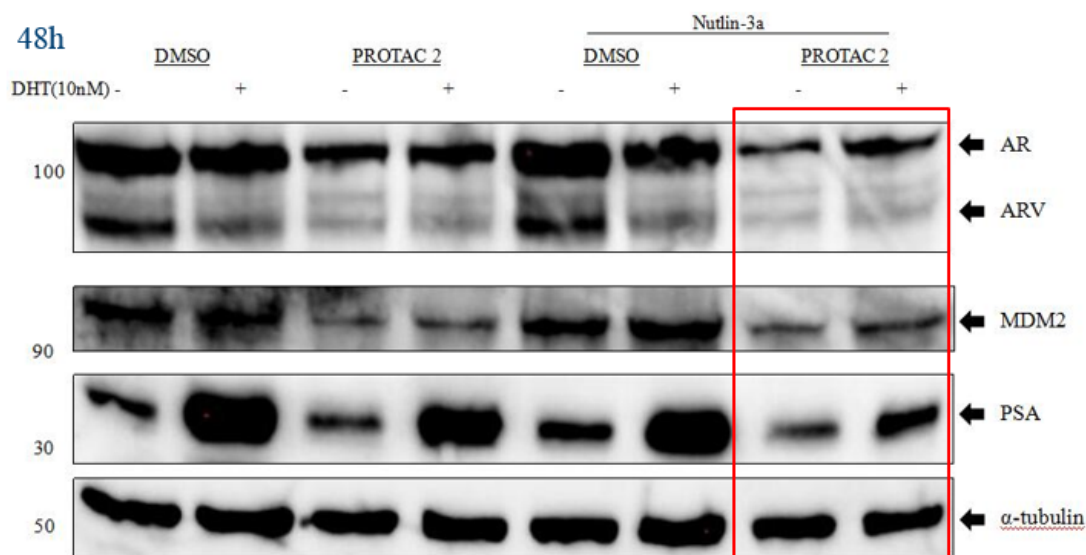


Figure 163; Addition of PROTAC 2 and Nutlin-3a to VCaP cell

- AR and PSA levels are reduced in two PROTAC 2 groups and also in the combination group, but no remarkable difference is observed.

## 22.4 Cell proliferation assays

To test whether any of our PROTACs could influence the proliferation of PCa cell lines, cell proliferation assays were carried. The results are summarized below:

### 22.4.1 PROTAC 1 cell proliferation assay

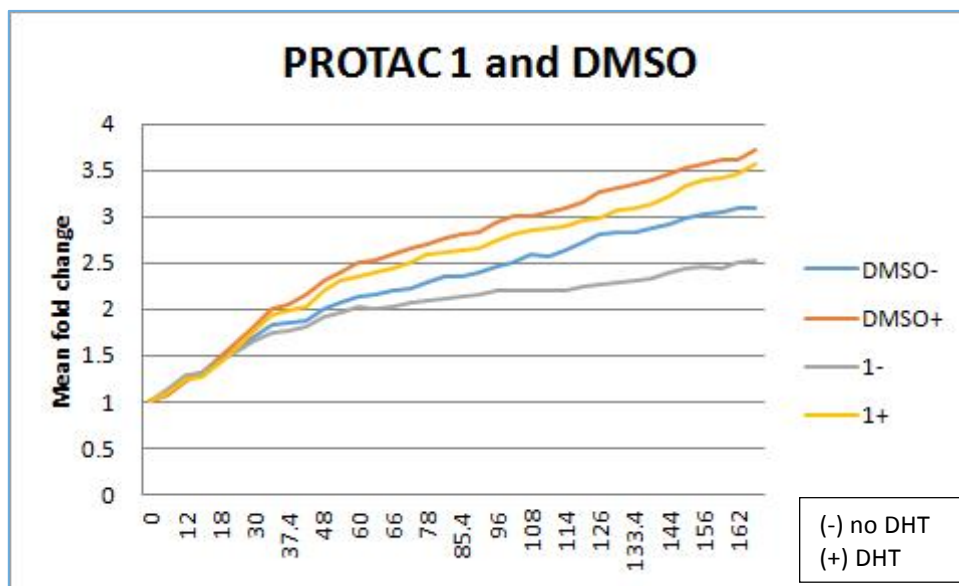


Figure 164; Cell proliferation assay in LNCaP cells for PROTAC 1 (x-axis = cell number)

- The proliferation is reduced upon PROTAC 1 addition, with or without DHT addition.

### 22.4.2 PROTAC 2 cell proliferation assay

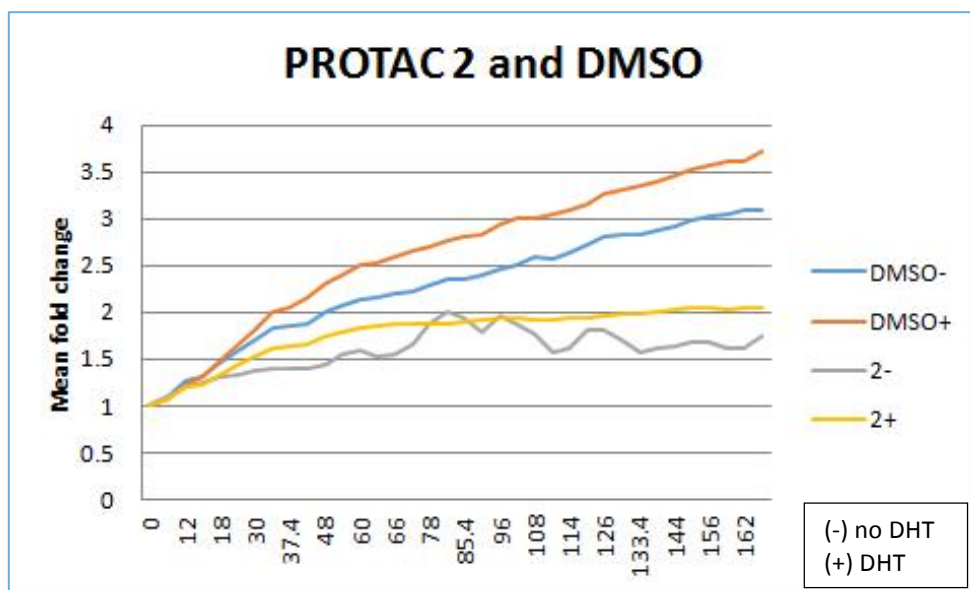


Figure 165; Cell proliferation assay in LNCaP cells for PROTAC 2 (x-axis = cell number)

- The proliferation is reduced upon PROTAC 2 addition, with or without DHT addition.

### 22.4.3 PROTAC 3 cell proliferation assay

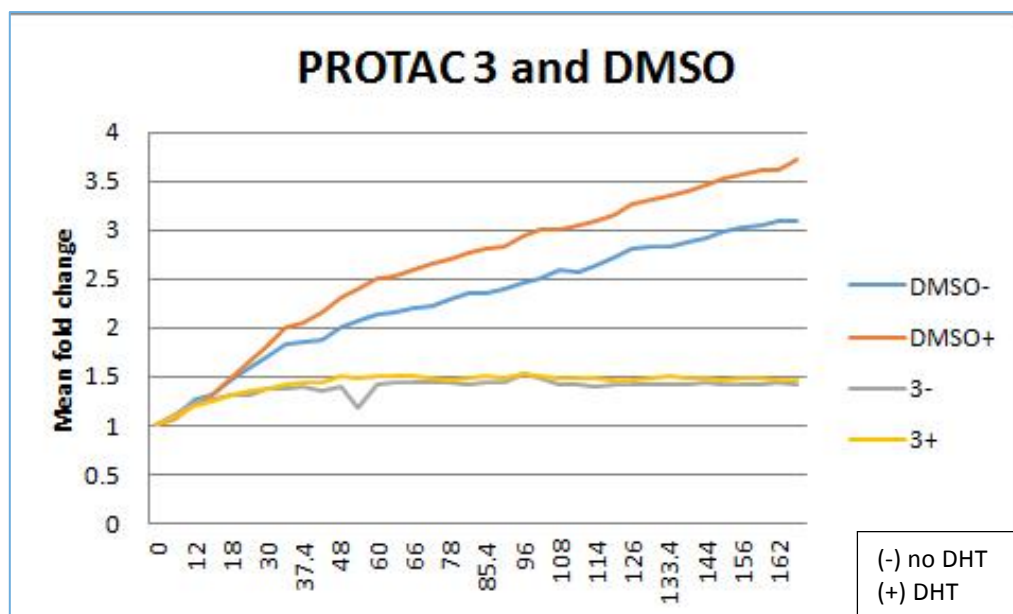


Figure 166; Cell proliferation assay in LNCaP cells for PROTAC 3 (x-axis = cell number)

- The proliferation is reduced upon PROTAC 3 addition, with or without DHT addition.

### 22.4.4 PROTAC 4 cell proliferation assay

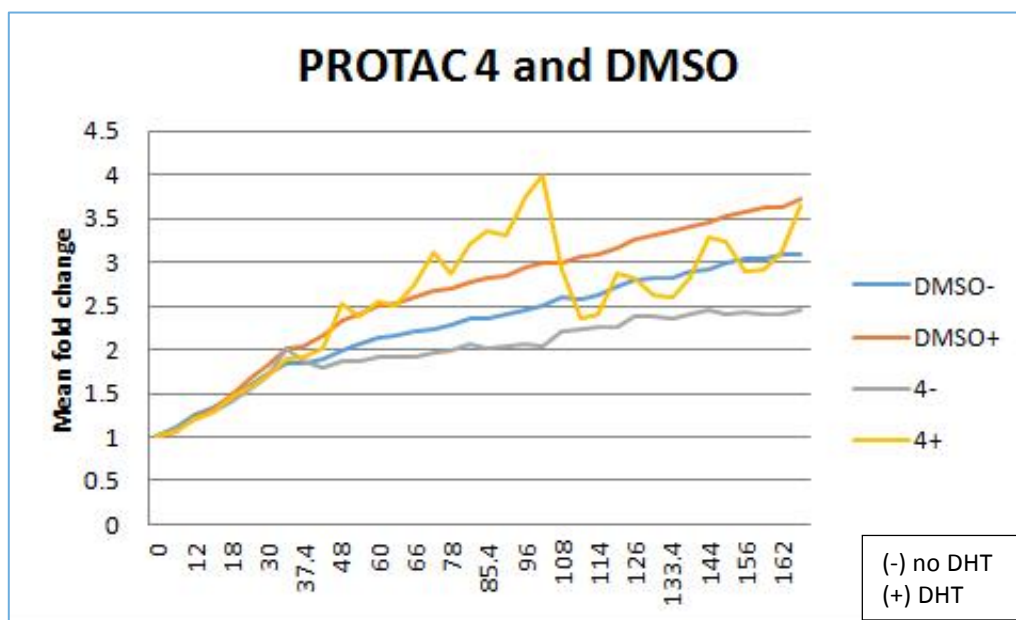


Figure 167; Cell proliferation assay in LNCaP cells for PROTAC 4 (x-axis = cell number)

- The proliferation is reduced upon PROTAC 5 addition, with or without DHT addition.



#### 22.4.5 PROTAC 5 cell proliferation assay

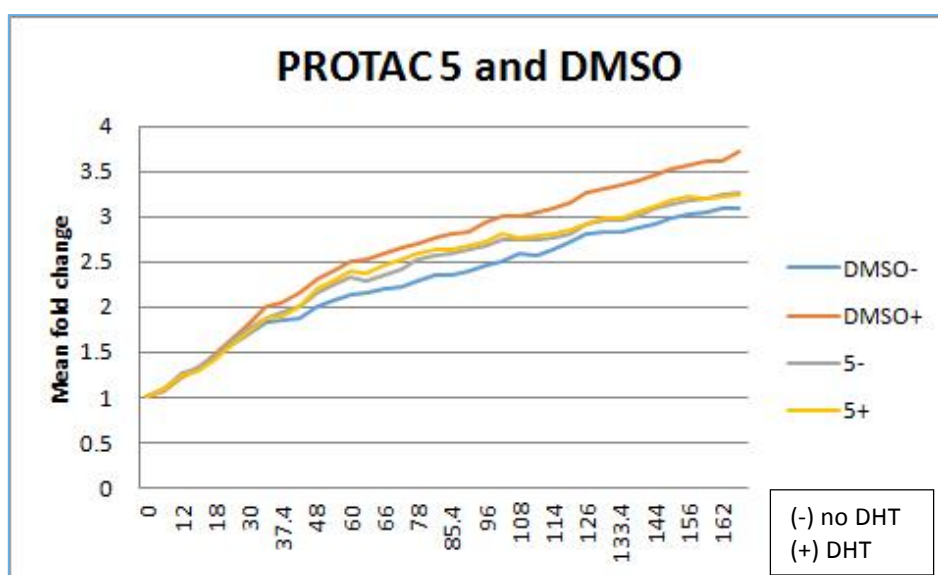


Figure 168; Cell proliferation assay in LNCaP cells for PROTAC 5 (**x-axys** = cell number)

As a general conclusion, all the PROTACs reduced the proliferation of LNCaP cells compared to the control groups. Among them, PROTAC 2 and 3 had the greater effects in LNCaP cells proliferation.



## Chapter 23: Conclusions and Future Work

The results of the biological assays were not conclusive. Further experimentation is required to assess if the cell proliferation assays were slightly successful because of induced degradation rather than simple inhibition, or cytotoxic effects. Addition of a proteasome inhibitor should be used to block the degradation of androgen receptors.

The fact that a cereblon-based PROTAC was relatively active compared to MDM2-based PROTACs confirms the tendency found in recent literature. Using thalidomide or VHL-based E3 ligase ligands could give us better chances of synthesizing successful PROTACs. Also, the optimisation of the linker (nature and length), and the nature of the moiety tethering the E3 ligase ligand and the linker should be considered.

The hook effect proved to be elusive and a proteasome inhibitor strategy to confirm degradation should be implemented to avoid inconclusive results. For this strategy, the addition of the proteasome inhibitor would tell us if the lower expression levels of androgen receptors are the consequence of inhibition by the flutamide molecule contained in our PROTACs rather than degradation. If the 26S proteasome is inhibited the levels of androgen receptor expression should be not affected by the PROTAC.

The enzymatic process to efficiently ubiquitinate a protein with the subsequent degradation can be harnessed and used to our purpose but in practice a complex development of events must happen. The failure of any of these can ruin our strategy and among the different steps for degradation of the targeted protein things can go wrong in several of those:<sup>284</sup>

1. The agent that brings in close proximity the cell degradation machinery and the protein must access the cell and PROTACs are usually large molecules with poor permeability.
2. The population of ternary complex must be sufficient to allow subsequent steps. Thermodynamics of this complex are hard to predict. A low population would not enable enough degradation and overpopulation can lead to the hook effect.
3. For the transfer of ubiquitin to the targeted protein the ternary complex must have an appropriate conformation, so the transfer of ubiquitin is directed to an appropriate receptor site, usually a lysine, so the rate of ubiquitin transfer is faster than the lifetime of the ternary complex.

4. At the same time, the speed of ubiquitination must overcome the action of deubiquitinase enzymes.
5. The ubiquitin chain formed in the targeted protein must be recognizable and accessible for the proteasome for degradation to happen. Different chain lengths/topologies can result in different responses.
6. Finally, the induced degradation rate must overcome the *de novo* synthesis of the targeted protein in order to ensure that the levels of the protein are diminished by degradation by PROTACs inside the cell.

Future work should be focused in the exploration of new E3 ligase ligands to target the androgen receptor. The use of MDM2 based PROTACs has not been optimised and new “degron” moieties will be tested.

The use of different linker lengths has shown to be critical for the design of a successful PROTACs and should be considered for future explorations.

PROTACs are a very promising strategy for the development of new therapeutic agents. In the following decades we will see the introduction of these molecules in the market by pioneers like Arvinas. The possibility of overcoming resistance and the capacities of these molecules to act as chemical probes can give the scientific community new insights about the functions of proteins within cells and the possibility to re-explore already known warheads that could be still active despite of new mutations developed in the targeted protein.

However, the lack of a rational design for the linker and the choice of an appropriate E3 ligase ligand depending on the targeted protein are the current handicaps of this emerging approach. The advent of more sophisticated computational chemistry approaches for the design of highly complex entities (like PROTACs) could give medicinal chemists new tools and the key to a faster approach to develop these heterobifunctional molecules. In the meantime, large amounts of manpower for the synthesis of large libraries with different linkers and E3 ligase ligands analogues are still necessary. Having said that, strategies inspired by combinatorial chemistry are emerging as I write to try to tackle these problems.

The fact that Arvinas and C4-Therapeutics have several clinical trials ongoing should make us feel optimistic about the future of this strategy that mixes a deep understanding in biology and organic chemistry for the delight of medicinal chemists.

## **Chapter 24: Experimental and Analysis**

### **24.1 Safety**

All procedures were carried out in line with the School of Natural Sciences Safety Policy. COSHH risk assessments were completed before starting any practical work.

### **24.2 Solvents and Reagents**

Chemicals were purchased from Sigma-Aldrich, Alfa Aesar, Apollo Scientific and Fluorochem and used without further purification. SureSeal™ or AcroSeal™ bottles of anhydrous solvents were purchased from Sigma-Aldrich or Acros, respectively. Deuterated solvents used for the determination of NMR spectra were purchased from Sigma-Aldrich.

### **24.3 Chromatography and Equipment**

Purifications by column chromatography were carried out using a Biotage SP4 automated flash system with UV monitoring at 298 nm and collection at 254 nm. Grace Resolv pre-packed flash cartridges were used in most cases for normal phase separations. Reveleris pre-packed NH silica cartridges and Reveleris C18 silica cartridges were also used where stated. Semi-preparative HPLC was carried out by Dr. Lauren Molyneux and Dr. Stephen Hobson using an Agilent 1200 HPLC system with a binary pump, autosampler, fraction collector and diode array detector, controlled by Agilent ChemStation software.

Where stated, reactions were carried out under microwave irradiation in sealed microwave vials with the use of a Biotage Initiator Sixty with a robotic sample bed. Reactions were irradiated at 2.45 GHz, and were able to reach temperatures between 60 and 250 °C. Heating was at a rate of 2-5 °C/s and the pressure was able to reach 20 bar.

### **24.4 Analytical Techniques**

Melting points were measured using a Stuart Scientific SMP3 apparatus. FTIR spectra were measured using a Agilent Cary 630 FTIR as a neat sample. UV spectra were recorded on a Hitachi U-2800A spectrophotometer and were performed in ethanol. HRMS were provided by the ESPRC National Mass Spectrometry Service, University of Wales, Swansea or Q-TOF in house.

LC-MS analyses were conducted using a Waters Acquity UPLC system with PDA and ELSD. When a 2 min gradient was used, the sample was eluted on Acquity UPLC BEH C18, 1.7 $\mu$ m, 2.1 x 50mm, with a flow rate of 0.6 ml/min using 5-95% 0.1% HCOOH in MeCN. Analytical purity of compounds was determined using Waters XTerra RP18, 5  $\mu$ m (4.6 x 150 mm) column at 1 ml/min using either 0.1% aq. ammonia and MeCN or 0.1% aq. HCOOH and MeCN with a gradient of 5-100% over 15 min.

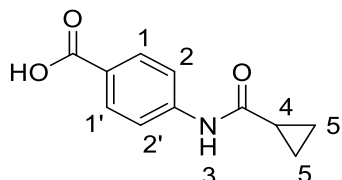
<sup>1</sup>H NMR spectra were obtained using a Bruker Avance III 500 spectrometer using a frequency of 500 MHz. <sup>13</sup>C and <sup>19</sup>F NMR spectra were acquired using the Bruker Avance III 500 spectrometer operating at a frequency of 125 MHz, and 470 MHz, respectively. The abbreviations for spin multiplicity are as follows: s = singlet; d = doublet; t = triplet; q = quartet, quin = quintet, sept = septet and m = multiplet. Combinations of these abbreviations are employed to describe more complex splitting patterns (e.g. dd = doublet of doublets) and where broadening of the peak is observed, spin multiplicity is accompanied by the prefix br = broad.

## Chapter 25: General Experimental

### 25.1 Synthesis of (15) and (17) and derivatives:

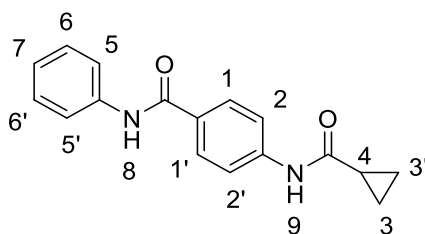
#### General procedure A for (14), (15), (16), (17), (18), (19) and (20)

##### 4-(Cyclopropanecarboxamido)benzoic acid (13)



To a suspension of *para*-aminobenzoic acid (2 g, 14.6 mmol) and  $\text{Na}_2\text{CO}_3$  (3.1 g, 29.2 mmol) in anhydrous THF (15 mL) was added dropwise cyclopropane carbonyl chloride (1.3 mL, 14.6 mmol) under  $\text{N}_2$  atmosphere at room temperature, the mixture was stirred for 16 hours, then evaporated *in vacuo* to give a pale yellow oil. The crude product was diluted in water and the pH adjusted to 2 after adding 2M HCl. The precipitate formed was filtered through a sintered funnel and dried in a vacuum oven for 16 hours affording a pale white solid (1.23 g, 44%). mp 320-321°C; UV  $\lambda_{\text{max}}$  (EtOH)/nm 273.0; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3281, 3015, 2667, 2541, 1677, 1650, 1523, 1421; LCMS (ESI)<sup>-</sup>  $m/z$  = 189.1 [ $\text{M} - \text{H}_2\text{O}$ ]<sup>-</sup>; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  12.67 (1H, s, -CO<sub>2</sub>H), 10.51 (1H, s, -NH-3), 7.88 (2H, d,  $J$  = 8.7 Hz, H-1, H-1'), 7.70 (2H, d,  $J$  = 8.7 Hz, H-2, H-2'), 1.82 (1H, p,  $J$  = 6.2 Hz, H-4), 0.83 (4H, d,  $J$  = 6.2 Hz, -CH<sub>2</sub>-5, -CH<sub>2</sub>-5'); <sup>13</sup>C NMR (126 MHz, DMSO)  $\delta$  172.6, 167.4, 143.8, 130.9 (x2), 125.3, 118.6 (x2), 15.2, 8.0 (x2); HRMS calc. for C<sub>11</sub>H<sub>11</sub>NO<sub>3</sub> [ $\text{M} - \text{OH}$ ]: 189.0790, found 189.0788

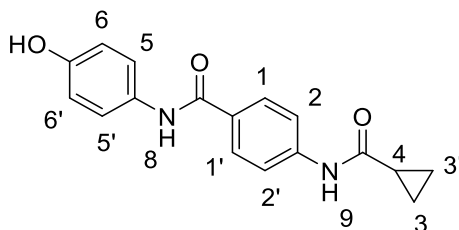
##### 4-(Cyclopropanecarboxamido)-*N*-phenylbenzamide (14)



To a stirred solution of 4-(cyclopropanecarboxamido)benzoic acid (0.17 g, 0.8 mmol), HATU (0.38 g, 1 mmol) and DIPEA (0.35 mL, 2.5 mmol) in DMF (5 mL) was added dropwise aniline (0.08 mL, 0.9 mmol) under  $\text{N}_2$  atmosphere at room temperature, the

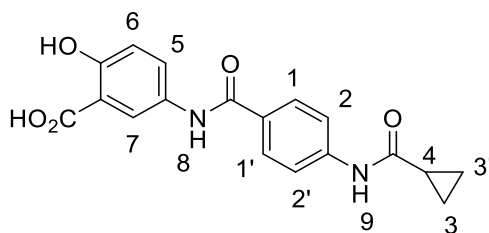
mixture was stirred for 16 hours, then evaporated *in vacuo*. The residue was taken up with EtOAc (20 mL), washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* to afford a yellow oil. Chromatography (SP4, EtOAc in Petrol, eluting 5-40%) gave a pale white solid (105 mg, 61%). mp 280-281°C; LCMS (ESI)<sup>-</sup> *m/z* = 279.0 [M - H]<sup>-</sup>; UV λ<sub>max</sub> (EtOH)/nm 286.2; IR ν<sub>max</sub>/cm<sup>-1</sup> 3275, 1645, 1596, 1517, 1438, 1406; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 10.49 (s, 1H, H-8), 10.11 (s, 1H, H-9), 7.93 (d, *J* = 8.8 Hz, 2H, H-1, H-1'), 7.77 (dd, *J* = 8.5, 1.0 Hz, 2H, H-5, H-5'), 7.73 (d, *J* = 8.8 Hz, 2H, H-2, H-2'), 7.35 (dd, *J* = 8.5, 7.5 Hz, 2H, H-6, H-6'), 7.09 (tt, *J* = 7.5 Hz, 1.0 Hz, 1H, H-7), 1.87 – 1.77 (m, 1H), 0.89 – 0.77 (m, 4H); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 172.5, 165.3, 142.7, 139.8, 129.4, 129.1 (x2), 129.0 (x2), 123.9, 120.8 (x2), 118.6 (x2), 15.2, 8.0 (x2); Analytical HPLC purity: 100% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>17</sub>H<sub>16</sub>N<sub>2</sub>O<sub>2</sub> [M+H]<sup>+</sup> 281.1285, found 281.1200

#### 4-(Cyclopropanecarboxamido)-*N*-(4-hydroxyphenyl)benzamide (16)



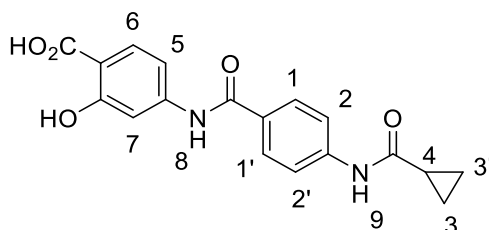
To a stirred solution of 4-(cyclopropanecarboxamido)benzoic acid (0.2 g, 0.97 mmol), HATU (0.45 g, 1.2 mmol) and DIPEA (0.41 mL, 2.9 mmol) in DMF (5 mL) was added portionwise 4-aminophenol (0.128 g, 1.2 mmol) under N<sub>2</sub> atmosphere at room temperature, the mixture was stirred for 16 hours, then evaporated *in vacuo*. The residue was taken up with EtOAc (20 mL), washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* to afford a yellow oil. Chromatography (SP4, C<sub>18</sub>, CH<sub>3</sub>CN [0.1% formic acid] in water [0.1% formic acid]; eluting 20-80%) gave a white solid (5 mg, 2%). R<sub>f</sub> 0.39 (C<sub>18</sub>, CH<sub>3</sub>CN [0.1% formic acid] in water [0.1% formic acid] (45:55)); mp 292-293°C; LCMS (ESI)<sup>-</sup> *m/z* = 295.1 [M - H]<sup>-</sup>; UV λ<sub>max</sub> (EtOH)/nm 287.4; IR ν<sub>max</sub>/cm<sup>-1</sup> 3289, 1638, 1609, 1510, 1434, 1408; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 10.49 (1H, s, H-8), 9.89 (1H, s, H-9), 9.28 (1H, br s, -OH), 7.90 (2H, d, *J* = 8.6 Hz, H-2, H-2'), 7.71 (2H, d, *J* = 8.6 Hz, H-1, H-1'), 7.51 (2H, d, *J* = 8.8 Hz, H-5, H-5'), 6.73 (2H, d, *J* = 8.8 Hz, H-6, H-6'), 1.82 (1H, p, *J* = 6.5 Hz, H-4), 0.92 – 0.69 (4H, m, -CH<sub>2</sub>-3, -CH<sub>2</sub>-3'); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 172.5, 164.8, 154.1, 142.5, 131.3, 129.6, 128.9 (x2), 122.7 (x2), 118.6 (x2), 115.4 (x2), 15.1, 7.9 (x2); Analytical HPLC purity: 99.6% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>17</sub>H<sub>16</sub>N<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup> 297.1234, found 297.1157

### 5-(4-(Cyclopropanecarboxamido)benzamido)-2-hydroxybenzoic acid (15)



**General procedure A:** Using 4-(cyclopropanecarboxamido)benzoic acid (0.2 g, 0.97 mmol) affording a pale white solid (190 mg, 57%). mp 300-301°C; LCMS (ESI)<sup>-</sup>  $m/z$  = 339.2 [M - H]<sup>-</sup>; UV  $\lambda_{\max}$  (EtOH)/nm 286.2; IR  $\nu_{\max}$ /cm<sup>-1</sup> 3264, 1652, 1607, 1523, 1449, 1408; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  10.49 (1H, s, H-8), 10.10 (1H, s, H-9), 8.26 (1H, d,  $J$  = 2.7 Hz, H-7), 7.93 (2H, d,  $J$  = 8.8 Hz, H-2, H-2'), 7.87 (1H, dd,  $J$  = 8.9, 2.7 Hz, H-5), 7.72 (2H, d,  $J$  = 8.8 Hz, H-1, H-1'), 6.95 (1H, d,  $J$  = 8.9 Hz, H-6), 1.82 (1H, p,  $J$  = 5.6 Hz, H-4), 0.89 – 0.74 (4H, m, -CH<sub>2</sub>-3, -CH<sub>2</sub>-3'); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  172.5, 172.2, 165.1, 157.8, 142.7, 131.4, 129.2, 129.0, 128.97 (x2), 122.2, 118.6, 117.4, 113.0, 15.1, 7.9 (x2); Analytical HPLC purity: 98.4% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>18</sub>H<sub>16</sub>N<sub>2</sub>O<sub>5</sub> [M+H]<sup>+</sup> 341.1132, found 341.1093

### 4-(4-(Cyclopropanecarboxamido)benzamido)-2-hydroxybenzoic acid (17)

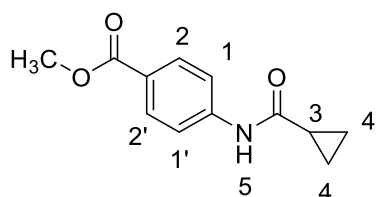


**General procedure A:** Using 4-(cyclopropanecarboxamido)benzoic acid (0.2 g, 0.97 mmol) affording a pale yellow oil. Chromatography (SP4, MeOH [0.1% formic acid] in DCM, eluting 1-15%) gave a pale white solid which was triturated in dry MeOH (5 mg, 2%). LCMS (ESI)<sup>-</sup>  $m/z$  = 339.2 [M - H]<sup>-</sup>; UV  $\lambda_{\max}$  (EtOH)/nm 291.8; IR  $\nu_{\max}$ /cm<sup>-1</sup> 3319, 1661, 1592, 1513, 1458; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  10.52 (1H, s, H-8), 10.31 (1H, s, H-9), 7.93 (2H, d,  $J$  = 8.8 Hz, H-2, H-2'), 7.74 (3H, d,  $J$  = 8.8 Hz, H-1, H-1', H-6), 7.52 (1H, d,  $J$  = 1.8 Hz, H-7), 7.32 (1H, dd,  $J$  = 8.7, 1.8 Hz, H-5), 1.82 (1H, p,  $J$  = 6.0 Hz, H-4), 0.89 – 0.79 (4H, m, -CH<sub>2</sub>-3, -CH<sub>2</sub>-3'); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  172.6, 165.8, 162.7, 162.5, 143.0, 131.2, 129.3 (x2), 128.9, 118.6 (x2), 111.4, 107.3, 15.2, 8.0 (x2); Analytical HPLC purity: 98.7% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>18</sub>H<sub>16</sub>N<sub>2</sub>O<sub>5</sub> [M+H]<sup>+</sup> 341.1132,



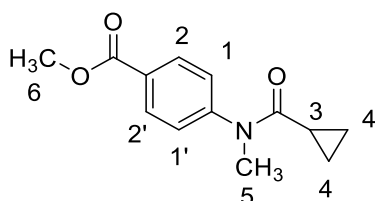
found 341.1066

### Methyl 4-(cyclopropanecarboxamido)benzoate (21)



To a stirred solution of 4-methyl-aminobenzoate (3 g, 20 mmol) and DIPEA (5.6 mL, 40 mmol) in DCM (20 mL) was added dropwise cyclopropanecarbonyl chloride (1.8 mL, 20 mmol) at 0°C under N<sub>2</sub> atmosphere, the mixture was warmed slowly to room temperature and stirred for 16 hours. The mixture was treated with saturated aqueous solution of NaHCO<sub>3</sub> and extracted with DCM (x3), the combined organic layers were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, C<sub>18</sub>, MeOH [0.1% NH<sub>3</sub>] in water, eluting 25-95%) gave a colourless oil (1.1 g, 25%). UV  $\lambda_{\text{max}}$  (EtOH)/nm 275.0; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3303, 3008, 2949, 2423, 2364, 1713, 1649, 1601, 1539, 1511, 1452, 1409; LCMS (ESI)<sup>-</sup>  $m/z$  = 218.2 [M - H]<sup>-</sup>; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  7.95 (2H, d,  $J$  = 8.9 Hz, H-2, H-2'), 7.68 (2H, d,  $J$  = 8.9 Hz, H-1, H-1'), 3.88 (3H, s, -CH<sub>3</sub>), 1.82 – 1.75 (1H, m, H-3), 1.00 – 0.93 (2H, m, -CH<sub>2</sub>-4), 0.91 – 0.85 (2H, m, -CH<sub>2</sub>-4'); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  173.8, 166.8, 143.3, 130.1 (x2), 124.7, 118.6 (x2), 51.0, 14.3, 6.9 (x2); HRMS calc. for C<sub>12</sub>H<sub>13</sub>NO<sub>3</sub> [M - MeO]<sup>-</sup> 189.0970, found 189.0701

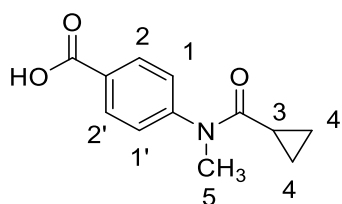
### Methyl 4-(*N*-methylcyclopropanecarboxamido)benzoate (22)



To a stirred solution of methyl 4-(cyclopropanecarboxamido)benzoate (1 g, 4.9 mmol) in anhydrous THF (20 mL) was added portionwise (60%) NaH (0.25 g, 6.2 mmol) at 0°C under N<sub>2</sub> atmosphere. The mixture stirred for 30 minutes and then iodomethane (0.38 mL, 6.2 mmol) was added dropwise. The mixture was allowed to warm to room temperature slowly and was stirred for 16 hours. The mixture was treated with saturated aqueous solution of NaHCO<sub>3</sub>, extracted with DCM (x3) and the organic layers were combined, washed with

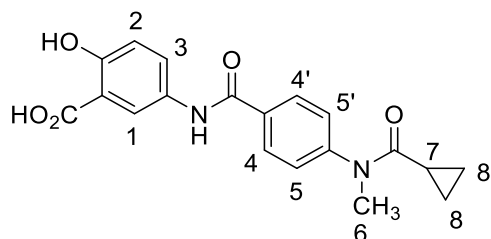
brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* to afford a yellow oil. Chromatography (SP4, C<sub>18</sub>, MeOH [0.1% NH<sub>3</sub>] in water [0.1% NH<sub>3</sub>], eluting 20-80%) affording a pale yellow oil (1.0 g, 95%). UV  $\lambda_{\text{max}}$  (EtOH)/nm 261.8; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3008, 2951, 1717, 1651, 1601, 1430; LCMS (ESI)<sup>-</sup>  $m/z$  = 232.3 [M - H]<sup>-</sup>; <sup>1</sup>H NMR (500 MHz, Chloroform-*d*)  $\delta$  8.09 (2H, d,  $J$  = 8.4 Hz, H-1, H-1'), 7.36 (2H, d,  $J$  = 8.4 Hz, H-2, H-2'), 3.93 (3H, s, -CH<sub>3</sub>-6), 3.33 (3H, s, -CH<sub>3</sub>-5), 1.41 (1H, tt,  $J$  = 8.0, 4.6 Hz, H-3), 1.05 (2H, dt,  $J$  = 6.9, 3.3 Hz, -CH<sub>2</sub>-4), 0.67 (2H, dq,  $J$  = 6.9, 3.7 Hz, -CH<sub>2</sub>-4'); <sup>13</sup>C NMR (125 MHz, Chloroform-*d*)  $\delta$  173.4, 166.3, 148.4, 130.9 (x2), 128.8, 127.0 (x2), 52.3, 37.3, 12.9, 8.8 (x2)

#### 4-(*N*-Methylcyclopropanecarboxamido)benzoic acid (23)



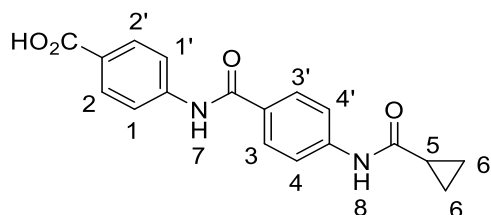
To a stirred solution of methyl 4-(*N*-methylcyclopropanecarboxamido)benzoate (1 g, 4.3 mmol) in anhydrous THF (10 mL) was added 2M NaOH in water (10 mL). The mixture was stirred at room temperature for 16 hours, 2M HCl was added to pH=2 and extracted with EtOAc (x3), the combined organic extracts were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a pale white solid. Chromatography (SP4, C<sub>18</sub>, MeOH [0.1% NH<sub>3</sub>] in water [0.1% NH<sub>3</sub>] eluting 5-95%) to give a colourless oil (0.94 g, 99.99%). LCMS (ESI)<sup>-</sup>  $m/z$  = 203.2 [M - H<sub>2</sub>O]<sup>-</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm 272.5; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3018, 2931, 1737, 1716, 1686, 1665, 1645, 1619, 1493, 1482; LCMS (ESI)<sup>-</sup>  $m/z$  = 203.2 [M - H<sub>2</sub>O]<sup>-</sup>; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  8.13 (2H, d,  $J$  = 8.3 Hz, H-1, H-1'), 7.50 (2H, d,  $J$  = 8.3 Hz, H-2, H-2'), 3.94 (3H, s, -CH<sub>3</sub>-5), 1.56 – 1.43 (1H, m, H-3), 0.88 – 1.03 (2H, m, H-4), 0.84 – 0.68 (2H, m, H-4'); <sup>13</sup>C NMR (125 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  175.6, 167.6, 149.5, 132.0 (x2), 130.4, 128.4 (x2), 52.8, 37.8, 13.7, 9.2 (x2); HRMS (ESI)<sup>+</sup> [M+H]<sup>+</sup>  $m/z$  Calc. for C<sub>12</sub>H<sub>13</sub>NO<sub>3</sub> [M - OH] 203.0946, found 203.0940

#### 2-Hydroxy-5-(4-(*N*-methylcyclopropanecarboxamido)benzamido)benzoic acid (24)



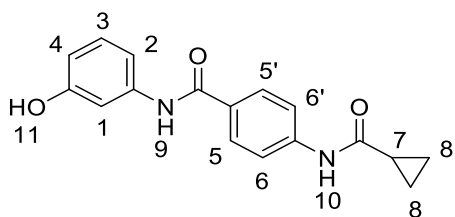
**General procedure A:** Using 4-(*N*-methylcyclopropanecarboxamido)benzoic acid (0.25 g, 1.2 mmoles) affording a pale yellow oil. Chromatography (SP4, MeOH [0.1% NH<sub>3</sub>] in water [0.1% NH<sub>3</sub>], eluting 5-40%) gave a pale white solid (5 mg, 1%). mp 238-239°C; LCMS (ESI)<sup>-</sup>  $m/z$  = 336.9 [M – H – H<sub>2</sub>O]<sup>-</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm 289.2; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3264, 1673, 1635, 1488, 1434, 1392, 1280, 1214, 1123, 1010; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  8.23 (1H, d,  $J$  = 2.6 Hz, H-1), 8.05 (2H, d,  $J$  = 8.4 Hz, H-5, H-5'), 7.76 (1H, dd,  $J$  = 8.9, 2.6 Hz, H-3), 7.51 (2H, d,  $J$  = 8.4 Hz, H-4, H-4'), 6.95 (1H, d,  $J$  = 8.9 Hz, H-2), 3.33 (3H, br s, -CH<sub>3</sub>-6), 1.56 – 1.44 (1H, m, H-7), 0.99 – 0.93 (2H, m, -CH<sub>2</sub>-8), 0.78 – 0.69 (2H, m, -CH<sub>2</sub>-8'); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  175.7, 173.2, 167.7, 160.3, 148.3, 135.2, 131.2, 130.6, 130.2 (x2), 128.5, 124.4, 118.4 (x2), 113.7, 37.9, 13.7, 9.1 (x2); Analytical HPLC purity: 99.2% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>19</sub>H<sub>18</sub>N<sub>2</sub>O<sub>5</sub> [M+H]<sup>+</sup> 355.1289, found 355.1221

#### 4-(4-(Cyclopropanecarboxamido)benzamido)benzoic acid (18)



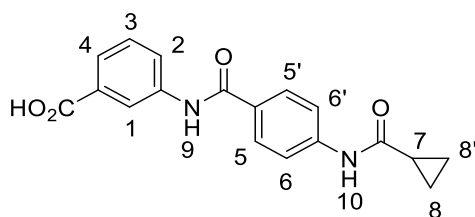
**General procedure A:** Using 4-(cyclopropanecarboxamido)benzoic acid (0.2 g, 0.97 mmoles) affording a yellow oil. Chromatography (SP4, C<sub>18</sub>, CH<sub>3</sub>CN [0.1% formic acid] in water [0.1% formic acid]; eluting 5-60%) gave a pale white solid (2 mg, 1%). mp 238-239°C; LCMS (ESI)<sup>-</sup>  $m/z$  = 322.7 [M – H – H<sub>2</sub>O]<sup>-</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm 291.8; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3269, 1646, 1515, 1408; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  12.74 (1H, s, -CO<sub>2</sub>H), 10.52 (1H, s, -NH-7), 10.39 (1H, s, -NH-8), 7.95 (2H, d,  $J$  = 8.7 Hz, -H-4, H-4'), 7.92 (4H, m, H-1, H-1', H-2, H-2'), 7.75 (2H, d,  $J$  = 8.7 Hz, H-3, H-3'), 1.83 (1H, p,  $J$  = 6.0 Hz, H-5), 0.94 – 0.78 (4H, m, -CH<sub>2</sub>-6, -CH<sub>2</sub>-6'); <sup>13</sup>C NMR (126 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  130.6 (x2), 129.3 (x2), 119.8 (x2), 118.6 (x2), 15.2, 8.0 (x2) (7 missing carbons); Analytical HPLC purity: 100% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>18</sub>H<sub>16</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup> 325.1183, found 325.1111

#### 4-(Cyclopropanecarboxamido)-N-(3-hydroxyphenyl)benzamide (19)



To a stirred solution of 4-(cyclopropanecarboxamido)benzoic acid (0.18 g, 0.88 mmol), HATU (0.4 g, 1.1 mmol) and DIPEA (0.37 mL, 2.6 mmol) in anhydrous DMF (5 mL) was added portionwise 3-aminophenol (0.12 g, 1.1 mmol) under N<sub>2</sub> atmosphere at room temperature. The mixture was stirred for 16 hours and then evaporated *in vacuo*. The residue was taken up with EtOAc (20 mL), washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, C<sub>18</sub>, MeOH [0.1% formic acid] in water [0.1% formic acid]; eluting 20-80%) gave a pale white solid. (5 mg, 2%). mp 250-251°C; UV  $\lambda_{\text{max}}$  (EtOH)/nm 285.0; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3535, 3364, 3279, 1640, 1606, 1518, 1444, 1407, 1308, 1185; LCMS (ESI)<sup>-</sup>  $m/z$  = 295.3 [M - H]<sup>-</sup>; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  10.48 (1H, s, H-9), 9.97 (1H, s, H-10), 9.38 (1H, s, -OH-11), 7.91 (2H, d,  $J$  = 8.6 Hz, H-6, H-6'), 7.72 (2H, d,  $J$  = 8.6 Hz, H-5, H-5'), 7.35 (1H, br s, H-1), 7.15 (1H, d,  $J$  = 7.9 Hz, H-2), 7.10 (1H, t,  $J$  = 7.9 Hz, H-3), 6.49 (1H, d,  $J$  = 7.9 Hz, H-4), 1.82 (1H, p,  $J$  = 6.6 Hz, H-7), 0.90 – 0.77 (4H, m, -CH<sub>2</sub>-8, -CH<sub>2</sub>-8'); <sup>13</sup>C NMR (126 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  172.5, 165.3, 157.9, 142.6, 140.8, 129.6, 129.5, 129.1 (x2), 118.6 (x2), 111.5, 111.1, 107.9, 15.1, 7.9 (x2); Analytical HPLC purity: 95.1% (CH<sub>3</sub>CN [0.1% NH<sub>3</sub>]); HRMS calc. for C<sub>17</sub>H<sub>16</sub>N<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup> 297.1234, found 297.1380

#### 3-(4-(Cyclopropanecarboxamido)benzamido)benzoic acid (20)

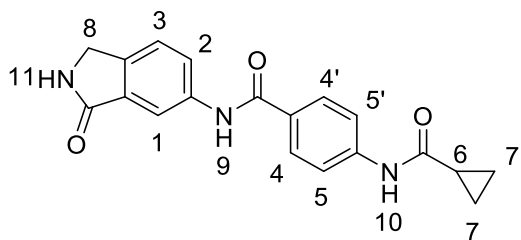


**General procedure A:** Using 4-(cyclopropanecarboxamido)benzoic acid (0.4 g, 2.0 mmol) affording a pale yellow oil. Chromatography (SP4, C<sub>18</sub>, CH<sub>3</sub>CN [0.1% formic acid] in water [0.1% formic acid]; eluting 5-55%) gave a pale white solid (21 mg, 3%). mp: 318-319°C; UV  $\lambda_{\text{max}}$  (EtOH)/nm 285.0, 508.6; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3278, 1683, 1640, 1591, 1521, 1407; LCMS (ESI)<sup>-</sup>  $m/z$  = 307.3 [M - H<sub>2</sub>O - H]<sup>-</sup>; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  12.98 (1H, s, -CO<sub>2</sub>H),

10.50 (1H, s, H-9), 10.29 (1H, s, H-10), 8.42 (1H, s, H-1), 8.04 (1H, d,  $J = 7.9$  Hz, H-2), 7.96 (2H, d,  $J = 8.7$  Hz, H-6, H-6'), 7.74 (2H, d,  $J = 8.7$  Hz, H-5, H-5'), 7.67 (1H, d,  $J = 7.9$  Hz, H-4), 7.47 (1H, t,  $J = 7.9$  Hz, H-3), 1.94 - 1.73 (1H, m, H-7), 1.02 - 0.77 (4H, m, -CH<sub>2</sub>-8, -CH<sub>2</sub>-8'); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  172.6, 167.7, 165.5, 142.9, 140.0, 131.7, 129.3, 129.2, 129.0, 124.8, 124.7, 121.5, 118.6, 15.2, 8.0 (x2); Analytical HPLC purity: 98.1% CH<sub>3</sub>CN [0.1% NH<sub>3</sub>]; HRMS calc. for C<sub>18</sub>H<sub>16</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup> 325.1183, found 325.1111

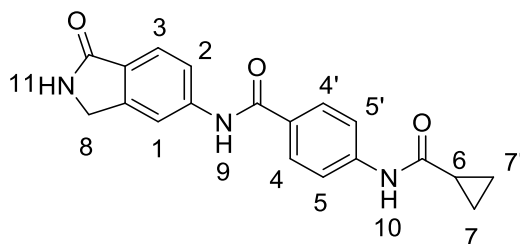
#### General procedure B for (27), (25) and (26)

##### 4-(Cyclopropanecarboxamido)-*N*-(3-oxoisindolin-5-yl)benzamide (27)



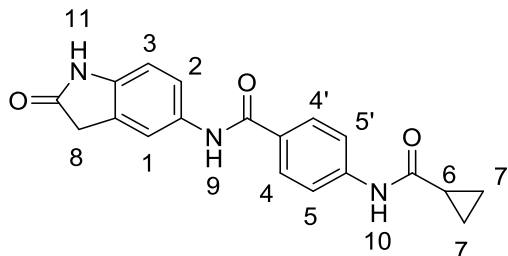
To a stirred solution of 4-(cyclopropanecarboxamido)benzoic acid (0.2 g, 0.97 mmol), HATU (0.45 g, 1.2 mmol) and DIPEA (0.41 mL, 2.9 mmol) in DMF (5 mL) was added portionwise 6-aminoisindolin-1-one (0.17 g, 1.2 mmol) under N<sub>2</sub> atmosphere at room temperature. The mixture was stirred for 16 hours, then filtered through a sintered funnel and washed with EtOAc (5 mL). The precipitate formed was collected and dried in a vacuum oven for 16 hours affording a white solid (160 mg, 49%). mp 366-367°C; LCMS (ESI)<sup>-</sup>  $m/z$  = 334.3 [M - H]<sup>-</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm 285.2; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3255, 3188, 3070, 1664, 1598, 1523, 1451, 1407; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  10.55 (1H, s, H-9), 10.32 (1H, s, H-10), 8.58 (1H, s, H-11), 8.18 (1H, d,  $J = 1.8$  Hz, H-1), 7.96 (2H, d,  $J = 8.9$  Hz, H-5, H-5'), 7.94 (1H, dd,  $J = 8.3, 1.8$  Hz, H-2), 7.75 (2H, d,  $J = 8.9$  Hz, H-4, H-4'), 7.54 (1H, d,  $J = 8.3$  Hz, H-3), 4.34 (2H, s, -CH<sub>2</sub>-8), 1.92 - 1.79 (1H, m, H-6), 0.96 - 0.72 (4H, m, -CH<sub>2</sub>-7, -CH<sub>2</sub>-7'); <sup>13</sup>C NMR (126 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  172.6, 170.4, 165.5, 142.9, 139.5, 139.3, 133.5, 129.1 (x2), 124.2, 124.0, 118.6 (x2), 114.7, 45.1, 15.1, 8.0 (x2) (1 missing carbon); Analytical HPLC purity: 100% (CH<sub>3</sub>CN [0.1% NH<sub>3</sub>]); HRMS calc. for C<sub>18</sub>H<sub>16</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup> 336.1343, found 336.1281

#### 4-(Cyclopropanecarboxamido)-*N*-(1-oxoisindolin-5-yl)benzamide (25)

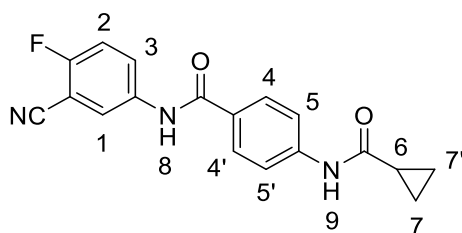


**General procedure B:** Using 4-(cyclopropanecarboxamido)benzoic acid (0.2 g, 0.97 mmol) affording a white solid (100 mg, 31%). mp 343-344°C; LCMS (ESI)<sup>+</sup>  $m/z$  = 336.1 [M + H]<sup>+</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm 297.6; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3263, 3201, 1652, 1605, 1522, 1432; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  10.51 (1H, s, H-9), 10.38 (1H, s, H-10), 8.41 (1H, s, H-11), 8.12 (1H, d,  $J$  = 1.7 Hz, H-1), 7.95 (2H, d,  $J$  = 8.8 Hz, H-5, H-5'), 7.79 (1H, dd,  $J$  = 8.3, 1.7 Hz, H-2), 7.75 (2H, d,  $J$  = 8.8 Hz, H-4, H-4'), 7.63 (d,  $J$  = 8.3 Hz, H-3), 4.37 (2H, s, -CH<sub>2</sub>-8), 1.93 – 1.76 (1H, m, H-6), 0.91 – 0.75 (4H, m, -CH<sub>2</sub>-7, -CH<sub>2</sub>-7'); <sup>13</sup>C NMR (126 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  172.6, 170.2, 165.7, 145.6, 142.9, 142.7, 129.2 (x2), 129.1, 128.0, 123.6, 120.1, 118.6 (x2), 115.0, 45.4, 15.2, 8.0 (x2); Analytical HPLC purity: 97.6% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>18</sub>H<sub>16</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup> 336.1343, found 336.1277

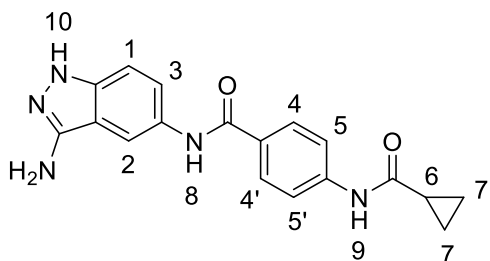
#### 4-(Cyclopropanecarboxamido)-*N*-(2-oxoisindolin-5-yl)benzamide (26)



**General procedure B:** Using 4-(cyclopropanecarboxamido)benzoic acid (0.2 g, 0.97 mmol) affording a white solid (169 mg, 52%). mp 337-338°C; LCMS (ESI)<sup>+</sup>  $m/z$  = 336.1 [M + H]<sup>+</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm 291.2; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3259, 1730, 1684, 1630, 1604, 1524, 1444, 1407; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  10.46 (1H, s, H-9), 10.32 (1H, s, H-10), 9.98 (1H, s, H-11), 7.91 (2H, d,  $J$  = 8.7 Hz, H-5, H-5'), 7.72 (2H, d,  $J$  = 8.7 Hz, H-4, H-4'), 7.66 (1H, d,  $J$  = 1.9 Hz, H-1), 7.51 (1H, dd,  $J$  = 8.4, 1.9 Hz, H-2), 6.78 (1H, d,  $J$  = 8.4 Hz, H-3), 3.50 (2H, s, -CH<sub>2</sub>-8), 1.87 – 1.77 (1H, m, H-6), 0.88 – 0.77 (4H, m, -CH<sub>2</sub>-7, -CH<sub>2</sub>-7'); <sup>13</sup>C NMR (126 MHz, DMSO)  $\delta$  176.8, 172.5, 165.0, 142.6, 140.1, 133.7, 129.5, 128.9 (x2), 126.3, 120.3, 118.6 (x2), 118.2, 109.2, 36.5, 15.1, 7.9 (x2); Analytical HPLC purity: 98.8% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>18</sub>H<sub>16</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup> 336.1343, found

***N*-(3-Cyano-4-fluorophenyl)-4-(cyclopropanecarboxamido)benzamide (28)**

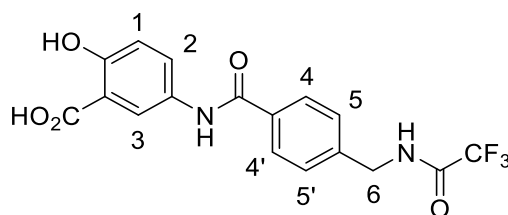
To a stirred solution of 4-(cyclopropanecarboxamido)benzoic acid (0.2 g, 0.97 mmol), HATU (0.45 g, 1.2 mmol) and DIPEA (0.41 mL, 2.9 mmol) in anhydrous DMF (5 mL) was added portionwise 5-amino-2-fluorobenzonitrile (0.13 g, 0.97 mmol) under N<sub>2</sub> atmosphere at room temperature. The mixture was stirred for 5 days, then diluted with EtOAc (20 mL) and washed with brine (x3), dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, EtOAc in Petrol, eluting 1-50%) to give a white solid (85 mg, 27%). *R*<sub>f</sub> 0.82 (silica, EtOAc in Petrol (40/60)); mp 297-298°C; LCMS (ESI)<sup>+</sup> *m/z* = 322.4 [M - H]<sup>-</sup>; UV λ<sub>max</sub> (EtOH)/nm 287.6; IR ν<sub>max</sub>/cm<sup>-1</sup> 3342, 3299, 2240, 1684, 1643, 1591, 1522; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 10.52 (1H, s, H-8), 10.45 (1H, s, H-9), 8.28 (1H, dd, *J* = 5.7, 2.6 Hz, H-1), 8.07 (1H, dq, *J* = 8.8, 2.6 Hz, H-3), 7.94 (2H, d, *J* = 8.6 Hz, H-5, H-5'), 7.75 (2H, d, *J* = 8.6 Hz, H-6, H-6'), 7.54 (t, *J* = 8.8 Hz, H-2), 1.83 (1H, p, *J* = 6.1 Hz, H-6), 0.92 – 0.77 (4H, m, -CH<sub>2</sub>-7, -CH<sub>2</sub>-7'); <sup>13</sup>C NMR (126 MHz, DMSO-*d*<sub>6</sub>) δ 172.6, 165.6, 159.7, 157.7, 143.1, 136.9, 129.2 (x2), 128.5, 128.04, 127.98, 124.5, 118.7 (x2), 117.5, 117.3, 114.5, 100.3, 100.2, 15.2, 8.0 (x2); <sup>19</sup>F NMR (471 MHz, DMSO-*d*<sub>6</sub>) δ -115.0; Analytical HPLC purity: 99.2% (CH<sub>3</sub>CN [0.1% NH<sub>3</sub>]); HRMS calc. for C<sub>18</sub>H<sub>14</sub>FN<sub>3</sub>O<sub>2</sub> [M+H]<sup>+</sup> 324.1143, found 324.1079

***N*-(3-Amino-1-*H*-indazol-5-yl)-4-(cyclopropanecarboxamido)benzamide (29)**

To a stirred solution of *N*-(3-cyano-4-fluorophenyl)-4-(cyclopropanecarboxamido)benzamide (0.65 g, 0.2 mmol) in dry EtOH (1.5 mL) and anhydrous THF (0.5 mL) was added

hydrazine monohydrate (12  $\mu$ L, 0.4 mmoles) under  $N_2$  atmosphere at room temperature. The mixture was stirred at 60°C for 24 hours. After observing starting material by LCMS hydrazine monohydrate (48  $\mu$ L, 3.2 mmoles) was added. The mixture was stirred at 60°C for 3 days, then filtered through a sintered funnel, washed with EtOH (5 mL) and the precipitate formed was collected and dried in a vacuum oven for 16 hours affording a pale brown solid (26 mg, 39%). mp 357-358°C; LCMS (ESI)<sup>+</sup>  $m/z$  = 334.1 [M - H]<sup>-</sup>; UV  $\lambda_{\max}$  (EtOH)/nm 270.2; IR  $\nu_{\max}/\text{cm}^{-1}$  3445, 3264, 1638, 1594, 1534, 1495, 1456, 1410; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  11.32 (1H, s, H-10), 10.47 (1H, s, H-8), 10.03 (1H, s, H-9), 8.08 (1H, d,  $J$  = 2.0 Hz, H-2), 7.95 (2H, d,  $J$  = 8.6 Hz, H-5, H-5'), 7.73 (2H, d,  $J$  = 8.6 Hz, H-4, H-4'), 7.42 (1H, dd,  $J$  = 8.9, 2.0 Hz, H-3), 7.21 (1H, d,  $J$  = 8.9 Hz, H-1), 5.27 (2H, s, -NH<sub>2</sub>), 1.83 (1H, p,  $J$  = 7.0 Hz, H-6), 0.94 – 0.75 (4H, m, -CH<sub>2</sub>-7, -CH<sub>2</sub>-7'); <sup>13</sup>C NMR (126 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  172.5, 165.0, 149.7, 142.5, 139.4, 129.9, 129.6, 128.9 (x2), 122.6, 118.6 (x2), 114.1, 112.9, 109.6, 15.1, 7.9 (x2); Analytical HPLC purity: 100% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>18</sub>H<sub>17</sub>N<sub>5</sub>O<sub>2</sub> [M+H]<sup>+</sup> 336.1455, found 336.1370

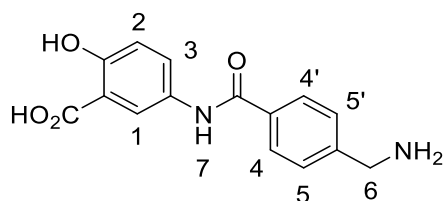
## 2-Hydroxy-5-(4-((2,2,2-trifluoroacetamido)methyl)benzamido)benzoic acid (32)



To a stirred solution of 4-(aminomethyl)benzoic acid (0.5 g, 3.3 mmoles) in dry DCM (15 mL) was added trifluoroacetic anhydride (1.2 mL, 6.6 mmoles) under  $N_2$  atmosphere at 0°C. The mixture was stirred for 30 minutes then anhydrous pyridine (0.11  $\mu$ L) was added dropwise under  $N_2$  atmosphere at 0°C and the mixture stirred for 2 hours, then evaporated *in vacuo*. The residue was dissolved in thionyl chloride (5 mL) and stirred at room temperature for 3 hours. The mixture was evaporated *in vacuo* and the residue dissolved in anhydrous THF (5 mL) and 5-amino-2-hydroxybenzoic acid (760 mg, 4.96 mmoles) in anhydrous THF (5 mL) was added dropwise at 0°C under  $N_2$  atm, stirred at room temperature for 16 hours, then diluted with saturated aqueous solution of NaHCO<sub>3</sub>, extracted with EtOAc (x3), the combined organic extracts were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a brown oil which was used in the next step without further purification (0.83 g, 65%).

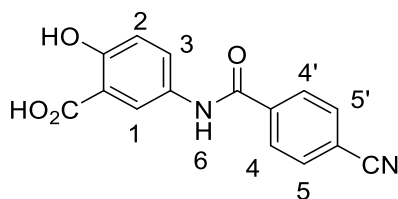


### 5-(4-(Aminomethyl)benzamido)-2-hydroxybenzoic acid (33)



To a stirred solution of 2-hydroxy-5-(4-((2,2,2-trifluoroacetamido)methyl)benzamido)benzoic acid (0.83 g, 2.17 mmol) in anhydrous dioxane (5 mL) was added dropwise 10 mL of a mixture  $\text{NH}_3/\text{H}_2\text{O}$  (1:9). The mixture was stirred at  $80^\circ\text{C}$  for 16 hours and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4,  $\text{C}_{18}$ ,  $\text{CH}_3\text{CN}$  [0.1%  $\text{NH}_3$ ] in water [0.1%  $\text{NH}_3$ ], eluting 5-75%) to give a pale black solid (5 mg, 1%).  $R_f$  0.61 ( $\text{C}_{18}$ ,  $\text{CH}_3\text{CN}$  [0.1%  $\text{NH}_3$ ] in water [0.1%  $\text{NH}_3$ ] (45:55)); mp  $277-278^\circ\text{C}$ ; IR  $\nu_{\text{max}}/\text{cm}^{-1}$  3393, 3254, 2116, 1630, 1532, 1477, 1428, 1375, 1340, 1312, 1239;  $^1\text{H}$  NMR (500 MHz,  $\text{DMSO}-d_6$ )  $\delta$  8.26 (1H, d,  $J = 2.6$  Hz, H-3), 8.12 (1H, s, H-7), 7.98 (2H, d,  $J = 8.2$  Hz, H-4, H-4'), 7.84 (1H, dd,  $J = 8.9, 2.6$  Hz, H-2), 7.58 (2H, d,  $J = 8.2$  Hz, H-5, H-5'), 6.98 (1H, d,  $J = 8.9$  Hz, H-1), 4.11 (2H, s,  $-\text{CH}_2-$ 6);  $^{13}\text{C}$  NMR (126 MHz,  $\text{DMSO}-d_6$ )  $\delta$  171.7, 165.3, 163.4, 159.2, 158.9, 158.6, 158.3, 157.7, 137.5, 135.0, 130.9, 129.3, 129.2 (x2), 128.3 (x2), 122.5, 117.6, 112.8, 42.2 (+1 drop of TFA for solubility); Analytical HPLC purity: 88.5% ( $\text{CH}_3\text{CN}$  [0.1% formic acid]); HRMS calc. for  $\text{C}_{15}\text{H}_{14}\text{N}_2\text{O}_4$  286.0954, found 286.1178

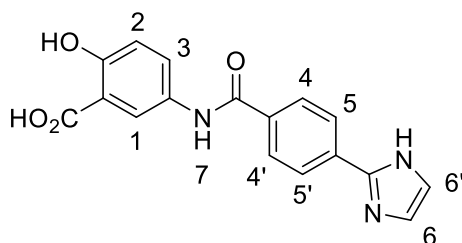
### 5-(4-Cyanobenzamido)-2-hydroxybenzoic acid (37)



A mixture of 4-cyanobenzoic acid (0.28 g, 1.9 mmol) in thionyl chloride (5 mL) and anhydrous DMF (catalytic) was stirred at  $80^\circ\text{C}$  under  $\text{N}_2$  atmosphere for 16 hours. The mixture was evaporated *in vacuo* and the residue redissolved in anhydrous THF (5 mL), then 4-amino-2-hydroxybenzoic acid (0.29 g, 1.9 mmol) in anhydrous THF (5 mL) was added dropwise and subsequently triethylamine (1.33 mL, 9.5 mmol). The mixture was stirred at room temperature for 48 hours, then treated with saturated aqueous solution of  $\text{NaHCO}_3$ , extracted with EtOAc (x3), the combined organic extracts were dried ( $\text{Na}_2\text{SO}_4$ ), filtered and evaporated *in vacuo* affording a pale yellow oil. Chromatography (SP4,  $\text{C}_{18}$ ,  $\text{CH}_3\text{CN}$  [0.1%

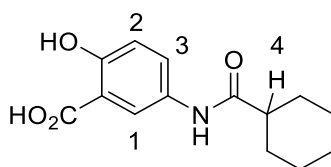
formic acid] in water [0.1% formic acid], eluting 5-55%) gave a white solid (75 mg, 14%).  $R_f$  0.76 ( $C_{18}$ ,  $CH_3CN$  [0.1% formic acid] in water [0.1% formic acid] (45:55)); mp 261-262°C; UV  $\lambda_{max}$  (EtOH)/nm 236.4, 318.2; IR  $\nu_{max}/cm^{-1}$  3388, 2232, 1982, 1668, 1619, 1549, 1492;  $^1H$  NMR (500 MHz,  $DMSO-d_6$ )  $\delta$  10.46 (1H, s, H-6), 8.27 (1H, d,  $J$  = 2.7 Hz, H-1), 8.11 (2H, d,  $J$  = 8.6 Hz, H-4, H-4'), 8.03 (2H, d,  $J$  = 8.6 Hz, H-5, H-5'), 7.88 (1H, dd,  $J$  = 8.9, 2.7 Hz, H-3), 6.99 (1H, d,  $J$  = 8.9 Hz, H-2);  $^{13}C$  NMR (126 MHz,  $DMSO$ )  $\delta$  172.0, 164.3, 158.1, 139.2, 132.9 (x2), 130.8, 129.0, 128.9 (x2), 122.5, 118.8, 117.6, 114.3, 113.1; Analytical HPLC purity: 97.8% ( $CH_3CN$  [0.1% formic acid]); HRMS calc. for  $C_{15}H_{11}N_2O_4$   $[M+H]^+$  283.0713, found 283.0991

### 5-(4-(1-*H*-Imidazol-2-yl)benzamido)-2-hydroxybenzoic acid (35)



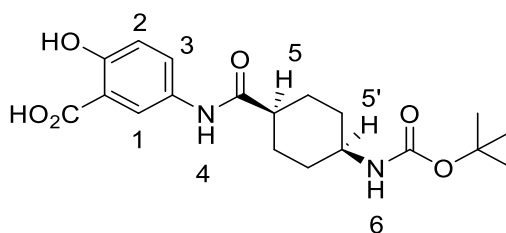
A mixture of 4-(1-*H*-Imidazol-2-yl)benzoic acid (0.18 g, 0.96 mmoles) in thionyl chloride (5 mL) and anhydrous DMF (catalytic) was stirred at 80°C under  $N_2$  atmosphere for 16 hours. The mixture was evaporated *in vacuo* and the residue redissolved in THF (5 mL), then 4-amino-2-hydroxybenzoic acid (0.15 g, 0.96 mmoles) in anhydrous THF (5 mL) was added dropwise. The mixture was stirred at room temperature for 16 hours, then was treated with saturated aqueous solution of  $NaHCO_3$ , extracted with EtOAc (x3), the combined organic extracts were dried ( $Na_2SO_4$ ), filtered and evaporated *in vacuo* affording a pale yellow oil. Chromatography (SP4,  $C_{18}$ ,  $CH_3CN$  [0.1% formic acid] in water [0.1% formic acid], eluting 5-50%), to give a white solid (5 mg, 2%).  $R_f$  0.79 ( $C_{18}$ ,  $CH_3CN$  [0.1% formic acid] in water [0.1% formic acid] (45:55)); mp 291-292°C; UV  $\lambda_{max}$  (EtOH)/nm 310.4, 210.0; IR  $\nu_{max}/cm^{-1}$  3280, 3158, 1631, 1538, 1429;  $^1H$  NMR (500 MHz,  $DMSO-d_6$ )  $\delta$  10.21 (1H, s, H-7), 8.24 (1H, d,  $J$  = 2.7 Hz, H-1), 8.11 – 8.03 (4H, m, H-4, H-4', H-5, H-5'), 7.86 (1H, dd,  $J$  = 8.9, 2.7 Hz, H-3), 7.26 (2H, br s, H-6, H-6'), 6.92 (1H, d,  $J$  = 8.9 Hz, H-2);  $^{13}C$  NMR (126 MHz,  $DMSO-d_6$ )  $\delta$  172.1, 165.0, 158.2, 134.4, 133.4, 130.8, 128.6 (x2), 128.5, 125.1 (x2), 122.5, 117.2; Analytical HPLC purity: 98.9% ( $CH_3CN$  [0.1% formic acid]); HRMS calc. for  $C_{17}H_{13}N_3O_4$  323.0906, found 323.0286

### 5-(Cyclohexanecarboxamido)-2-hydroxybenzoic acid (39)



A mixture of cyclohexane carboxylic acid (0.18 g, 1.4 mmol) in thionyl chloride (5 mL) and anhydrous DMF (catalytic) was stirred at 80°C under N<sub>2</sub> atmosphere for 16 hours. The mixture was evaporated *in vacuo* and the residue redissolved in anhydrous THF (5 mL), then 4-amino-2-hydroxybenzoic acid (0.194 g, 1.26 mmol) in anhydrous THF (5 mL) was added dropwise and the mixture was stirred at room temperature for 16 hours. Then the mixture was treated with saturated aqueous solution of NaHCO<sub>3</sub>, extracted with EtOAc (x3), the combined organic extracts were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a pale yellow solid which was triturated in EtOAc and Et<sub>2</sub>O to give a white solid (15 mg, 4%). mp 288-289°C; UV  $\lambda_{\text{max}}$  (EtOH)/nm 314.6, 254.4, 218.8, 206.2; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3308; 2927, 2852, 2455, 2007, 1648, 1580, 1534, 1491, 1444; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  7.77 (1H, d, *J* = 2.7 Hz, H-1), 7.59 (1H, dd, *J* = 8.7, 2.7 Hz, H-3), 6.76 (1H, d, *J* = 8.7 Hz, H-2), 2.34 (1H, tt, *J* = 11.3, 3.3 Hz, H-4), 1.94 – 1.80 (4H, m, -CH<sub>2</sub>-), 1.74 (2H, dd, *J* = 9.2, 3.3 Hz, H-CH<sub>2</sub>-), 1.54 (2H, qd, *J* = 12.3, 11.8, 2.5 Hz, -CH<sub>2</sub>-), 1.38 (2H, qt, *J* = 12.3, 2.5 Hz, -CH<sub>2</sub>-), 1.33 – 1.26 (2H, m, -CH<sub>2</sub>-); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  176.0, 158.1, 128.9, 126.1, 123.0, 118.8, 115.6, 45.6, 29.3 (x2), 25.5, 25.5 (x2); Analytical HPLC purity: 97.2% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>18</sub>H<sub>17</sub>N<sub>5</sub>O<sub>2</sub>: 263.1158, found 263.1159

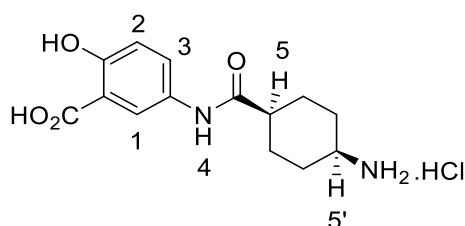
### 5-((1*S*,4*S*)-4-((*tert*-Butoxycarbonyl)amino)cyclohexane-1-carboxamido)-2-hydroxybenzoic acid (40)



To a stirred suspension of HBTU (0.47 g, 1.24 mmol) and DIPEA (0.45 mL, 2.47 mmol) in anhydrous DMF (2.5 mL) was added *tert*-butyl[*cis*-4-(anilino-carbonyl)cyclohexyl]carbamate (0.3 g, 1.24 mmol) at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred for 30 minutes and 5-amino-2-hydroxybenzoic acid (189 mg, 1.24 mmol) in anhydrous DMF (2.5 mL) was added and then mixture was stirred for

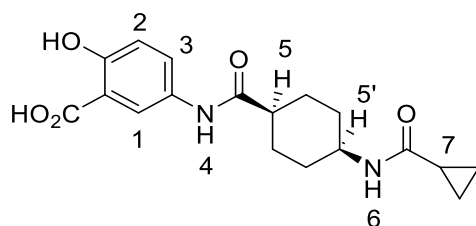
16 hours at room temperature. The mixture was treated with saturated aqueous solution of  $\text{NaHCO}_3$ , extracted with EtOAc (x3), the combined organic extracts were dried ( $\text{Na}_2\text{SO}_4$ ), filtered and evaporated *in vacuo* to give a brown oil which solidified upon standing and was used in the next step without further purification (385 mg, 82%). mp 221-222°C; UV  $\lambda_{\text{max}}$  (EtOH)/nm 283.4; IR  $\nu_{\text{max}}/\text{cm}^{-1}$  3625, 3404, 3300, 3238, 2981, 2936, 2865, 1657, 1524, 1490, 1435;  $^1\text{H}$  NMR (500 MHz, Methanol- $d_4$ )  $\delta$  9.47 (1H, s, -Ph-O-H), 8.49 – 8.28 (2H, br s, -CO<sub>2</sub>H, -NH-4), 7.91 (1H, d,  $J$  = 2.7 Hz, H-1), 7.51 (1H, dd,  $J$  = 8.7, 2.7 Hz, H-3), 6.75 (1H, br s, -NH-6), 6.68 (1H, d,  $J$  = 8.7 Hz, H-2), 3.53 – 3.45 (2H, m, H-5), 2.37 – 2.27 (1H, m, H-5'), 1.88 – 1.77 (2H, m, -CH<sub>2</sub>-), 1.74 – 1.67 (2H, m, -CH<sub>2</sub>-), 1.57 – 1.44 (4H, m, -CH<sub>2</sub>-(x2)), 1.39 (9H, s, -CH<sub>3</sub> (x3));  $^{13}\text{C}$  NMR (126 MHz, Methanol- $d_4$ )  $\delta$  173.7, 172.2, 158.3, 155.5, 130.0, 125.7, 122.0, 117.9, 116.2, 77.8, 49.1, 46.7 (x2), 28.8 (x2), 25.0, 19.3; Analytical HPLC purity: 98.6% ( $\text{CH}_3\text{CN}$  [0.1%  $\text{NH}_3$ ])

**5-((1*S*,4*S*)-4-Aminocyclohexane-1-carboxamido)-2-hydroxybenzoic acid hydrochloride (41)**



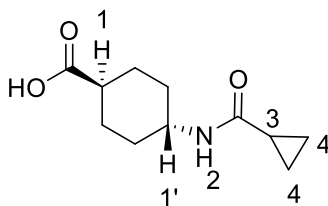
To a solution of 5-((1*S*,4*S*)-4-((tert-butoxycarbonyl)amino)cyclohexane-1-carboxamido)-2-hydroxy benzoic acid (0.38 g, 1 mmol) in anhydrous THF (5 mL) was added 2M HCl (10 mL) and the mixture was stirred at 80°C for 4 hours. The mixture was evaporated *in vacuo* to give a pale brown oil which was used in the next step without further purification (315 mg, 99.1%). mp 218-219°C; UV  $\lambda_{\text{max}}$  (EtOH)/nm 204.8, 254.6; IR  $\nu_{\text{max}}/\text{cm}^{-1}$  3241, 3053, 2923, 2885, 2654, 1646, 1616, 1524, 1486, 1422;  $^1\text{H}$  NMR (500 MHz, Methanol- $d_4$ )  $\delta$  8.23 (2H, br s, -NH-4), 8.14 (1H, d,  $J$  = 2.7 Hz, H-1), 8.03 (1H, br s, -NH-4), 7.96 (1H, d,  $J$  = 2.7 Hz, H-1), 7.63 (1H, dd,  $J$  = 8.9, 2.7 Hz, H-3), 7.58 (1H, dd,  $J$  = 8.9, 2.7 Hz, H-3), 7.10 (1H, d,  $J$  = 8.9 Hz, H-2), 6.90 (1H, d,  $J$  = 8.9 Hz, H-2), 3.64 – 3.13 (4H, m, -CH<sub>2</sub>-, H-5, H-5'), 2.96 – 2.60 (2H, m, -CH<sub>2</sub>-), 2.21 – 1.55 (8H, m, -CH<sub>2</sub>-(x2), -NH<sub>2</sub> (x2));  $^{13}\text{C}$  NMR (126 MHz, Methanol- $d_4$ )  $\delta$  176.4, 174.7, 171.8, 170.8, 161.8, 158.4, 129.9, 128.7, 124.7, 122.3, 121.5, 118.7, 116.9, 113.3, 112.1, 61.5, 40.5, 28.8, 26.9 (x2), 24.3 (x2); Analytical HPLC purity: 100% ( $\text{CH}_3\text{CN}$  [0.1%  $\text{NH}_3$ ]); HRMS (ESI<sup>+</sup>)  $[\text{M}+\text{H}]^+ m/z$  Calc. for  $\text{C}_{14}\text{H}_{19}\text{N}_2\text{O}_4$ : 279.1339, found 278.9222

**5-((1*S*,4*S*)-4-(Cyclopropanecarboxamido)cyclohexane-1-carboxamido)-2-hydroxybenzoic acid (42)**



Cyclopropane carbonyl chloride (0.27 mL, 3 mmol) was added dropwise to a solution of 5-((1*S*,4*S*)-4-aminocyclohexane-1-carboxamido)-2-hydroxybenzoic acid hydrochloride (0.315 g, 1 mmol) in anhydrous THF (5 mL) at room temperature. The mixture was stirred at room temperature for 48 hours and then was evaporated *in vacuo* affording a yellow oil which was triturated in MeOH to give a pale white solid (82 mg, 24%). mp 222-223°C; UV  $\lambda_{\text{max}}$  (EtOH)/nm 326.2, 254.2; IR  $\nu_{\text{max}}/\text{cm}^{-1}$  3266, 2937, 2098, 1723, 1645, 1544, 1485, 1445;  $^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ )  $\delta$  9.82 (1H, br s, H-4), 8.02 (1H, s, H-1), 7.78 (1H, br s, -NH-6), 7.63 (1H, d,  $J = 8.6$  Hz, H-3), 6.83 (1H, d,  $J = 8.6$  Hz, H-2), 4.13 (1H, t,  $J = 6.4$  Hz, H-5), 4.05 (1H, t,  $J = 6.4$  Hz, H-5) 3.66 (t,  $J = 6.4$  Hz, H-5'), 3.44 (1H, t,  $J = 6.4$  Hz, H-5'), 3.17 – 3.0 (1H, m, -CH<sub>2</sub>-), 2.68 – 2.56 (1H, m, -CH<sub>2</sub>-), 2.09 – 1.40 (6H, m, -CH<sub>2</sub>- (x3)), 0.90 - 0.78 (3H, m, -CH<sub>2</sub>-, H-7), 0.79 - 0.71 (2H, m, -CH<sub>2</sub>-);  $^{13}\text{C}$  NMR (126 MHz, DMSO- $d_6$ )  $\delta$  171.7, 157.5, 131.4, 127.3, 120.9, 117.3, 63.7, 45.5, 29.2, 27.4, 26.2, 24.6, 14.8 (x2), 7.4 (x3); Analytical HPLC purity: 97.9% (CH<sub>3</sub>CN [0.1% formic acid])

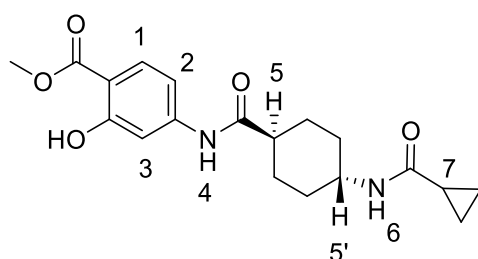
**(1*S*,4*R*)-4-(Cyclopropanecarboxamido)cyclohexane-1-carboxylic acid (44)**



To a stirred suspension of HBTU (0.87 g, 2.3 mmol) and triethylamine (0.16 mL, 2.30 mmol) in anhydrous THF (5 mL) was added (1*r*,4*r*)-4-aminocyclohexane-1-carboxylic acid (0.3 g, 2.1 mmol) at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred for 30 minutes and then cyclopropane carboxylic acid (198 mg, 1.15 mmol) in anhydrous THF (2.5 mL) was added. The mixture was stirred for 6 hours at room temperature and then was treated with water, extracted with EtOAc (x3), the combined organic extracts were dried

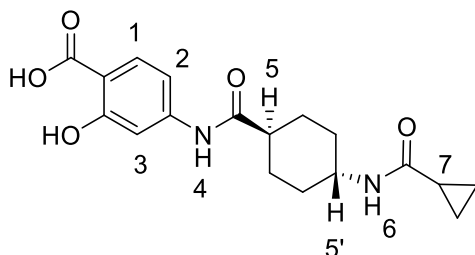
(Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a white solid which was triturated in EtOAc to give a white solid which was used in the next step without further purification (100 mg, 23%). <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 12.10 (1H, s, -CO<sub>2</sub>H), 7.96 (1H, d, *J* = 7.8 Hz, -NH-4), 3.46 (1H, dddt, *J* = 14.5, 10.8, 6.8, 3.4 Hz, H-1'), 2.14 (1H, tt, *J* = 11.9, 3.5 Hz, H-1), 1.93 – 1.85 (2H, m, -CH<sub>2</sub>-), 1.80 (2H, dd, *J* = 12.6, 2.9 Hz, -CH<sub>2</sub>-), 1.50 (2H, ddd, *J* = 12.6, 7.7, 4.9 Hz, H-3), 1.33 (2H, qd, *J* = 13.2, 3.1 Hz, -CH<sub>2</sub>-), 1.17 (2H, qd, *J* = 12.8, 3.3 Hz, -CH<sub>2</sub>-), 0.68 – 0.56 (4H, m, -CH<sub>2</sub>-6); <sup>13</sup>C NMR (126 MHz, DMSO-*d*<sub>6</sub>) δ 179.2, 171.1, 53.3, 48.7, 28.5 (x2), 22.8 (x2), 15.1, 9.1 (x2)

**Methyl-4-((1*S*,4*R*)-4-(cyclopropanecarboxamido)cyclohexane-1-carboxamido)-2-hydroxybenzoate (45)**



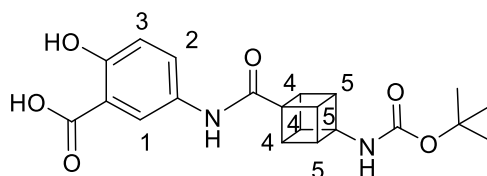
To a solution of EDCI.HCl (227 mg, 1.2 mmol) and DMAP (174 mg, 1.4 mmol) in anhydrous DMF (5 mL) at room temperature under N<sub>2</sub> atmosphere was added methyl 5-amino-2-hydroxybenzoate (79 mg, 0.47 mmol) in anhydrous DMF (2.5 mL). The mixture was stirred at room temperature for 48 hours and then was evaporated *in vacuo* to give a brown solid which was purified through and SCX-2 cartridge and washed with MeOH affording a white solid which was used in the next step without further purification (88 mg, 52%). <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 7.81 (1H, d, *J* = 8.6 Hz, H-1), 6.68 (1H, dd, *J* = 8.6, 2.2 Hz, H-2), 6.45 (1H, d, *J* = 2.2 Hz, H-3), 3.78 (3H, s, -CH<sub>3</sub>), 3.75 – 3.61 (1H, m, H-5), 2.58 (1H, m, H-5'), 2.29 – 2.17 (2H, m, -CH<sub>2</sub>-), 2.09 – 1.99 (2H, m, -CH<sub>2</sub>-), 1.71 – 1.60 (2H, m, -CH<sub>2</sub>-), 1.57 (1H, tt, *J* = 7.9, 4.7 Hz, H-7), 1.43 – 1.27 (2H, m, -CH<sub>2</sub>-), 0.91 – 0.82 (2H, m, -CH<sub>2</sub>-), 0.80 – 0.69 (2H, m, -CH<sub>2</sub>-)

**4-((1*S*,4*R*)-4-(Cyclopropanecarboxamido)cyclohexane-1-carboxamido)-2-hydroxybenzoic acid (46)**



2M NaOH (10 mL) was added to a stirred solution of methyl-4-((1*r*,4*r*)-4-(cyclopropanecarboxamido)cyclohexane-1-carboxamido)-2-hydroxybenzoate (88 mg, 0.24 mmoles) in anhydrous THF (2 mL). The mixture was stirred at 80°C for 5 hours and then was evaporated *in vacuo* affording a white solid. Chromatography (SP4, C<sub>18</sub>, CH<sub>3</sub>CN [0.1% formic acid] in water [0.1% formic acid]; eluting 40- 95%) gave a white solid (78 mg, 92%). mp 232-233 °C; UV  $\lambda_{\text{max}}$  (EtOH)/nm 298.8, 264.6; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 2935, 2857, 1554, 1418; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  7.59 (1H, d, *J* = 8.3 Hz, H-3), 6.14 (1H, dd, *J* = 8.4, 2.2 Hz, H-2), 6.09 (1H, d, *J* = 2.2 Hz, H-1), 3.73 – 3.58 (1H, m), 2.08 (1H, t, *J* = 12.2 Hz), 1.96 (4H, t, *J* = 16.8 Hz), 1.60 – 1.43 (3H, m), 1.26 (2H, q, *J* = 12.9, 12.0 Hz), 0.83 (2H, dt, *J* = 4.5, 3.1 Hz, -CH<sub>2</sub>), 0.72 (2H, dq, *J* = 7.2, 3.7 Hz, -CH<sub>2</sub>) (+2 CH<sub>3</sub>CN); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  183.5, 179.1, 174.0, 162.6, 152.6, 131.4, 108.7, 105.6, 100.2, 46.3, 32.1 (x2), 29.0 (x2), 22.8, 13.5, 5.7 (x2) (+CH<sub>3</sub>CN); Analytical HPLC purity: 98.6% (CH<sub>3</sub>CN [0.1% formic acid])

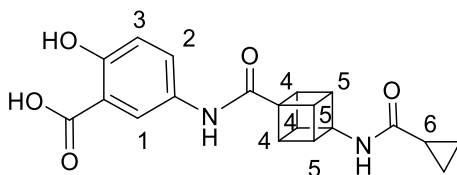
**5-((1*s*,2*R*,3*s*,4*r*,5*S*,6*r*,7*R*,8*S*)-4-((*tert*-butoxycarbonyl)amino)cubane-1-carboxamido)-2-hydroxybenzoic acid (48)**



To a stirred solution of HATU (70 mg, 0.18 mmoles) and TEA (50  $\mu$ L, 0.36 mmoles) in anhydrous DMF (2.5 mL) was added (1*s*,2*R*,3*s*,4*r*,5*S*,6*r*,7*R*,8*S*)-4-((*tert*-butoxycarbonyl)amino)cubane-1-carboxylic acid (50 mg, 0.19 mmoles) at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred for 30 minutes and 5-amino-2-hydroxybenzoic acid (10 mg, 0.07 mmoles) in anhydrous DMF (2.5 mL) was added and then mixture was stirred for 16 hours at room temperature. The mixture was treated with saturated aqueous

solution of NaHCO<sub>3</sub>, extracted with EtOAc (x3), the combined organic extracts were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* to give a brown oil. Chromatography (SP4, C<sub>18</sub>, MeOH [0.1% formic acid] in DCM [0.1% formic acid]; eluting 1- 10%) gave a yellow oil (15 mg, 20%). <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 7.81 (1H, s, H-1 ), 7.45 (1H, dd, *J* = 8.8, 2.6 Hz, H-2), 6.67 (1H, d, *J* = 8.7 Hz, H-3), 4.04 (3H, s, -CH-4 (x3)), 3.94 (3H, s, -CH-5 (x3)), 1.36 (9H, s, -CH<sub>3</sub> (x3))

**5-((1s,2R,3s,4r,5S,6r,7R,8S)-4-(cyclopropanecarboxamido)cubane-1-carboxamido)-2-hydroxybenzoic acid (50)**



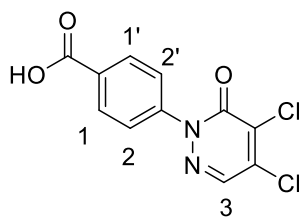
To a stirred solution of 5-((1s,2R,3s,4r,5S,6r,7R,8S)-4-((tert-butoxycarbonyl)amino)cubane-1-carboxamido)-2-hydroxybenzoic acid (15 mg, 0.04 mmoles) in dry DCM (2 mL) was added TFA (0.2 mL) at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred for 4 hours and then evaporated *in vacuo*. The residue was dissolved in dry THF (2 mL) and trimethylamine (3 μL, 0.19 mmoles) and cyclopropanecarbonyl chloride (1 μL, 0.04 mmoles) was added dropwise at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred for 24 hours at room temperature. The mixture was treated with water, extracted with EtOAc (x3), the combined organic extracts were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* to give a yellow oil. Chromatography (SP4, C<sub>18</sub>, MeOH [0.1% formic acid] in DCM [0.1% formic acid]; eluting 20- 80%) gave a yellow oil (2.5 mg, 18%). <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 7.98 (1H, d, *J* = 2.3 Hz, H-1), 7.56 (1H, dd, *J* = 8.9, 2.3 Hz, H-2), 6.78 (1H, d, *J* = 8.9 Hz, H-3), 4.18 – 4.06 (3H, m, -CH-4 (x3)), 4.04 – 3.90 (3H, m, -CH-5 (x3)), 1.51 (1H, m, -CH-6), 0.82 – 0.71 (2H, m, -CH<sub>2</sub>), 0.68 (2H, m, -CH<sub>2</sub>); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>) δ 174.6, 172.1, 158.6, 129.5, 128.5, 122.7, 116.6, 66.6, 58.3, 49.6, 44.9, 13.2, 6.1; Analytical HPLC purity: 94.8114% (CH<sub>3</sub>CN [0.1% formic acid])

## 25.2 Synthesis of (54) and derivatives:

### General procedure B for (54) and (55)

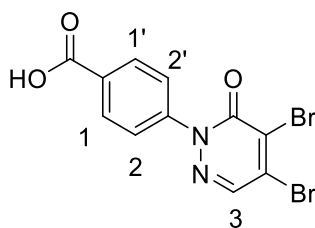


#### 4-(4,5-Dichloro-6-oxopyridazin-1(6-*H*)-yl)benzoic acid (54)



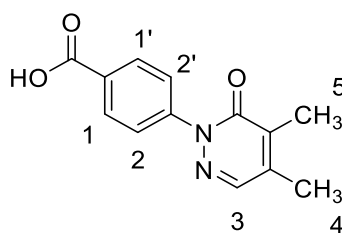
To 4-hydrazinobenzoic acid (0.5 g, 3.2 mmol) in a mixture of EtOH and H<sub>2</sub>O (1:1) (8 ml : 8 ml) was added mucochloric acid (0.6 g, 3.2 mmol) and 0.1 mL of 12M HCl (catalytic amount). The mixture was stirred at 65°C for 16 hours and then was filtered through a sintered funnel and washed with cold water affording 4-(4,5-dichloro-6-oxopyridazin-1(6-*H*)-yl)benzoic acid as a yellow solid (0.7 g, 76%). *R<sub>f</sub>* 0.63 (silica, EtOAc/petrol (1:1) + 1 drop of acetic acid); mp: 331-332 °C; LCMS (ESI<sup>-</sup>) *m/z* = 284.1 [M - H]<sup>-</sup>; UV λ<sub>max</sub> (EtOH)/nm 324; IR ν<sub>max</sub>/cm<sup>-1</sup>: 1690, 1653, 1572, 1428; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 13.23 (1H, br s, CO<sub>2</sub>H), 8.39 (1H, s, H-3), 8.09 (2H, d, *J* = 8.6 Hz, H-2, H-2'), 7.72 (2H, s, *J* = 8.6 Hz, H-1, H-1'); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 167.0, 156.1, 144.8, 137.3, 136.5, 134.6, 131.2, 130.3 (x2), 126.3 (x2); Analytical HPLC purity: 100% (CH<sub>3</sub>CN, [0.1% NH<sub>3</sub>]); HRMS calc. for C<sub>11</sub>H<sub>6</sub>Cl<sub>2</sub>N<sub>2</sub>O<sub>3</sub> [M-OH] 267.9806, found 267.9752

#### 4-(4,5-Dibromo-6-oxopyridazin-1(6-*H*)-yl)benzoic acid<sup>285</sup> (55)



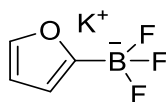
**General procedure B:** Using mucobromic acid (169 mg, 0.66 mmol) affording a pale yellow solid (105 mg, 58%). *R<sub>f</sub>* 0.68 (silica, EtOAc/Petrol (1:1) + HOAc (1%)); mp: 308-309°C; LCMS (ESI<sup>-</sup>) *m/z* = 372.9 [M - H]<sup>-</sup>; UV λ<sub>max</sub> (EtOH)/nm 286.5; IR ν<sub>max</sub>/cm<sup>-1</sup> 1691, 1650, 1561, 1427; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 13.21 (1H, br s, -COOH), 8.33 (1H, s, H-3), 8.07 (2H, d, *J* = 8.4 Hz, H-2, H-2'), 7.71 (2H, d, *J* = 8.4 Hz, H-1, H-1'); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 167.0, 156.3, 144.9, 138.9, 131.5, 131.4, 131.1, 130.3 (x2), 126.3 (x2); Analytical HPLC purity: 99.7% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>11</sub>H<sub>6</sub>Br<sub>2</sub>N<sub>2</sub>O<sub>3</sub> [M-OH] 357.8776, found 357.8721; (lit., LCMS (ESI<sup>-</sup>) *m/z* = 371, 373, 375 [M - H]<sup>-</sup>)

#### 4-(4,5-Dimethyl-6-oxopyridazin-1(6-*H*)-yl)benzoic acid (**57**)



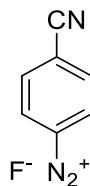
To a solution of (**55**) (0.2 g, 0.5 mmol) and Tetrakis(triphenylphosphine)palladium (31 mg, 0.03 mmol) in anhydrous THF (5 mL) was added (1.0 M in hexane)  $\text{Me}_3\text{Al}$  (4.3 mL, 4.3 mmol) dropwise at room temperature under  $\text{N}_2$  atmosphere. The mixture was stirred at reflux temperature for 1 hour and then was cooled to room temperature and quenched with 3 N HCl. The mixture was extracted with DCM (x3), the combined organic extracts were washed with brine, dried ( $\text{Na}_2\text{SO}_4$ ), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4,  $\text{C}_{18}$ ,  $\text{CH}_3\text{CN}$  [0.1% formic acid] in water [0.1% formic acid], eluting 5-50%) gave a pale yellow solid (18 mg, 98%).  $R_f$  0.37 ( $\text{C}_{18}$ ,  $\text{CH}_3\text{CN}$  [0.1% formic acid] in water [0.1% formic acid] (20/80)); mp: 297-298°C; LCMS (ESI $^-$ )  $m/z$  = 243.1 [ $\text{M} - \text{H}$ ] $^-$ ; UV  $\lambda_{\text{max}}$  (EtOH)/nm 257.1; IR  $\nu_{\text{max}}$ /cm $^{-1}$  1688, 1657, 1603, 1430;  $^1\text{H}$  NMR (500 MHz,  $\text{DMSO}-d_6$ )  $\delta$  13.13 (1H, br s, -COOH), 8.04 (2H, d,  $J$  = 8.4 Hz, H-1, H-1'), 7.94 (1H, s, H-3), 7.70 (2H, d,  $J$  = 8.4 Hz, H-2, H-2'), 2.22 (3H, s,  $\text{CH}_3$ -5), 2.09 (3H, s,  $\text{CH}_3$ -4);  $^{13}\text{C}$  NMR (125 MHz,  $\text{DMSO}-d_6$ )  $\delta$  140.7, 139.0, 137.0, 130.1 (x2), 125.9 (x2), 16.5, 12.6; Analytical HPLC purity: 97.9% ( $\text{CH}_3\text{CN}$  [0.1% formic acid]); HRMS calc. for  $\text{C}_{13}\text{H}_{12}\text{N}_2\text{O}_3$  [ $\text{M}-\text{OH}$ ] 228.0899, found 228.0841

#### Potassium trifluoro(furan-2-yl)borate (**61**)



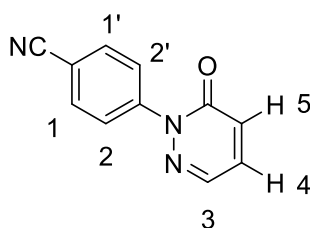
To a solution of 2-furanylboronic acid (1 g, 8.9 mmol) in MeOH (3 mL) at 0°C 2.1 g of  $\text{KHF}_2$  was added portionwise (2.1 g, 26.8 mmol) under  $\text{N}_2$  atmosphere. The mixture was stirred at 0°C for 1 hour and then the solvent was evaporated *in vacuo* affording potassium trifluoro(furan-2-yl)borate as a yellow solid (1.55 g, 99.99%) that was used in the next step without further purification.

#### 4-Cyanobenzenediazonium fluoride (62)



To a solution of 1.9 ml of boron trifluoride ethyl etherate (1.9 mL, 15 mmoles) in anhydrous DCM (20 mL) at 0°C was added 4-aminonitrile (1.2 g, 10 mmoles) portionwise under N<sub>2</sub> atmosphere. The mixture was stirred at 0°C for 20 minutes and then cooled down to -15°C for the addition of *tert*-BuONO (1.5 mL, 12 mmoles) dropwise. The mixture was stirred at -15°C under N<sub>2</sub> atmosphere for 1 hour and then the mixture was allowed to warm to room temperature slowly. The precipitate formed was filtered through a sintered funnel and washed with DCM. The filtrate was evaporated *in vacuo* and gave a pale yellow solid (1.51 g, 99.99%) which was used in the next step without further purification.

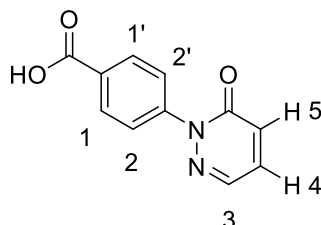
#### 4-(6-Oxopyridazin-1(6-*H*)-yl)benzonitrile<sup>286</sup> (63)



To a solution of potassium trifluoro(furan-2-yl)borate (1.55 g, 8.91 mmoles) in H<sub>2</sub>O (10 mL) was added 4-cyanobenzenediazonium fluoride (1.51 g, 10.13 mmoles) in H<sub>2</sub>O (10 mL) at 5°C dropwise. The mixture was allowed to warm to room temperature and stirred for 3 hours. The mixture was extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording an orange solid. Chromatography (SP4, silica, EtOAc/ Petrol, eluting 10-40%) gave a pale yellow solid (0.2 g, 80%). *R*<sub>f</sub> 0.28 (silica, EtOAc in Petrol (2:8)); mp 165-166 °C ; LCMS (ESI<sup>+</sup>) *m/z* = 198.2 [M + H]<sup>+</sup>; UV λ<sub>max</sub> (EtOH)/nm: 226.5; IR ν<sub>max</sub>/cm<sup>-1</sup> 1698, 1647, 1593, 1566, 1556; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 8.13 (1H, dd, *J* = 3.8, 1.6 Hz, H-4), 8.00 (2H, d, *J* = 8.7 Hz, H-1), 7.85 (2H, d, *J* = 8.7 Hz, H-2), 7.53 (1H, dd, *J* = 9.5, 3.8 Hz, H-3); 7.13 (1H, dd, *J* = 9.5, 1.6 Hz, H-5); <sup>13</sup>CNMR (Chloroform-*d*) δ 159.2, 144.3, 137.0, 132.2, 131.2, 131.1, 125.3, 117.9, 111.3;

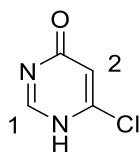
HRMS calc. for  $C_{11}H_7N_3O$   $[M+H]^+$  198.0662, found 198.0594; (lit., mp 158-159 °C; LCMS (ESI<sup>+</sup>)  $m/z$ = 198  $[M + H]^+$ )

#### 4-(6-Oxopyridazin-1(6H)-yl)benzoic acid (56)

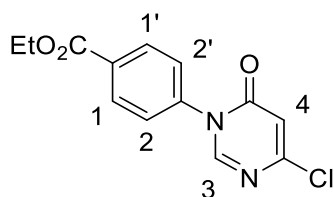


4-(6-Oxopyridazin-1(6H)-yl)benzonitrile (0.2 g, 1.1 mmol) was dissolved in a mixture of sulfuric acid (60 v/v) in water. The mixture was stirred at room temperature for 16 hours and then was extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried ( $Na_2SO_4$ ), filtered and the solvents evaporated *in vacuo* affording a yellow oil. Chromatography (SP4,  $C_{18}$ ,  $CH_3CN$  [0.1% formic acid] in water [0.1% formic acid], eluting 0-50%) gave a pale yellow solid (80 mg, 33%).  $R_f$  0.67 ( $CH_3CN$  [0.1% formic acid] in water [0.1% formic acid] (20/80)); mp 164-165 °C; LCMS (ESI<sup>+</sup>)  $m/z$   $[M + H]^+$  217.2; UV  $\lambda_{max}$  (EtOH)/nm: 287.8; IR  $\nu_{max}/cm^{-1}$  2231, 1661, 1590, 1507, 1415;  $^1H$  NMR (500 MHz, Methanol- $d_4$ )  $\delta$  8.10 (1H, dd,  $J$  = 3.8, 1.6 Hz, H-4), 7.93 – 7.84 (4H, m, H-1, H-1', H-2, H-2'), 7.51 (1H, dd,  $J$  = 9.5, 3.8 Hz, H-3), 7.12 (1H, dd,  $J$  = 9.5, 1.6 Hz);  $^{13}C$  NMR (125 MHz, Methanol- $d_4$ )  $\delta$  162.0, 146.5, 139.8, 134.0, 133.9 (x2), 132.0, 127.5 (x2), 119.1, 113.0; Analytical HPLC purity: 99.7%  $CH_3CN$  [0.1%  $NH_3$ ]; HRMS calc. for  $C_{11}H_8N_2O_3$   $[M+H]^+$  217.0608, found 217.0531

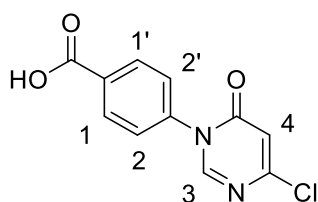
#### 6-Chloropyrimidin-4(1H)-one<sup>287</sup> (67)



To a stirred solution of 4,6-dichloropyrimidine (1 g, 6.7 mmol) in anhydrous dioxane (10 mL) was added a mixture of 12M HCl (1 mL) and water (9 mL). The mixture was stirred at 70 °C for 6 hours and then was concentrated *in vacuo* affording a pink oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 20-70%) gave a pale pink solid (0.5 g, 57%).  $R_f$  0.41 (EtOAc/Petrol (4:6)); mp 199-200 °C; LCMS (ESI<sup>+</sup>)  $[M + H]^+$  131.0;  $^1H$  NMR (500 MHz, Methanol- $d_4$ )  $\delta$  8.15 (1H, s, H-2), 6.54 (1H, s, H-1);  $^{13}C$  NMR (125 MHz, Methanol- $d_4$ )  $\delta$  164.2, 158.3, 151.4, 114.1; (lit., mp 193-194 °C)

**Ethyl 4-(4-chloro-6-oxopyrimidin-1(6-*H*)-yl)benzoate (68)**

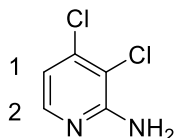
To a stirred solution of 6-chloropyrimidin-4(1H)-one (0.3 g, 1.9 mmol), 4-(ethoxy)benzene boronic acid (0.5 g, 2.3 mmol) and anhydrous pyridine (0.3 mL, 3.8 mmol) in anhydrous DCM (10 mL) was added copper (II) acetate (0.5 g, 2.7 mmol). The mixture was stirred at room temperature for 24 hours and then was filtered through celite and washed with EtOAc. The filtrate was evaporated *in vacuo* and the residue was taken up with EtOAc, washed with water and brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* to afford a brown oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 20-55%) gave a yellow oil (0.25 g, 62%). R<sub>f</sub> 0.29 (EtOAc/Petrol (1:9)); LCMS (ESI)<sup>+</sup> *m/z* = 278.0 [M + H]<sup>+</sup>; UV λ<sub>max</sub> (EtOH)/nm: 268.2; IR ν<sub>max</sub>/cm<sup>-1</sup> 1701, 1588, 1530, 1517, 1400; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 8.55 (1H, s, H-3), 8.12 (2H, d, *J* = 8.7 Hz, H-1, H-1'), 7.68 (2H, d, *J* = 8.7 Hz, H-2, H-2'), 6.79 (1H, s, H-4), 4.37 (2H, q, *J* = 7.1 Hz, -CH<sub>2</sub>-), 1.35 (3H, t, *J* = 7.1 Hz, -CH<sub>3</sub>); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 165.4, 159.7, 157.4, 152.4, 140.7, 131.1, 130.4 (x2), 128.1 (x2), 113.7, 61.6, 14.6; HRMS calc. for C<sub>13</sub>H<sub>11</sub>ClN<sub>2</sub>O<sub>3</sub> [M - EtO] 234.0196, found 234.0487

**Ethyl 4-(4-chloro-6-oxopyrimidin-1(6-*H*)-yl)benzoate (66)**

Ethyl 4-(4-chloro-6-oxopyrimidin-1(6H)-yl)benzoate (0.23 g, 0.7 mmol) was stirred in a mixture of H<sub>2</sub>SO<sub>4</sub>/H<sub>2</sub>O (1:5) at 80°C for 16 hours. The mixture was diluted with water (5 mL) and extracted with EtOAc (5 mL x6), the combined organic extracts were separated and evaporated *in vacuo* affording a colourless oil (20 mg, 11%). R<sub>f</sub> 0.46 (CH<sub>3</sub>CN [0.1% formic acid] in water [0.1% formic acid], eluting 20-80%); mp 307-308 °C; LCMS (ESI)<sup>+</sup> *m/z* = 251.1 [M + H]<sup>+</sup>; UV λ<sub>max</sub> (EtOH)/nm: 224.0; IR ν<sub>max</sub>/cm<sup>-1</sup> 3054, 2984, 1821, 2657, 2541, 2344, 2113, 1996, 1914, 1676, 1575, 1524, 1424; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 13.21 (1H, br s, -CO<sub>2</sub>H), 8.55 (1H, s, H-4), 8.09 (2H, d, *J* = 8.5 Hz, H-1, H-1'), 7.64 (2H, d, *J* = 8.5 Hz, H-2, H-2'), 6.79 (1H, s, H-3); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 167.0, 159.7, 157.4,

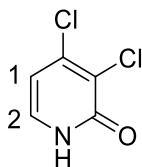
152.5, 140.4, 132.0, 130.6 (x2), 128.0 (x2), 113.7; Analytical HPLC purity: 100% (CH<sub>3</sub>CN [0.1% formic acid ]); HRMS calc. for C<sub>11</sub>H<sub>7</sub>ClN<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup> 251.0218, found 251.0179

### 3,4-Dichloropyridin-2-amine<sup>288</sup> (69)



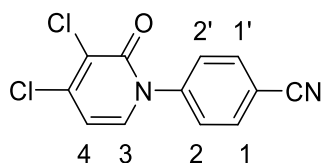
To a stirred solution of 2-amino-4-chloropyridine (2.5 g, 19.4 mmol) in conc. HCl (10 mL) was added H<sub>2</sub>O<sub>2</sub> (0.5 mL, 19.4 mmol) at 0°C. Then the mixture was stirred for 15 minutes at this temperature and then at room temperature for 2 hours. Then the mixture was slowly poured into a 2M aqueous solution of Na<sub>2</sub>CO<sub>3</sub> and the pH adjusted to 8, then the aqueous layer was extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 20-80%) gave a yellow oil (0.8 g, 27%). R<sub>f</sub> 0.25 (EtOAc/Petrol (1:4)); mp 105-106 °C; LCMS (ESI)<sup>+</sup> *m/z* = 164.0 [M + H]<sup>+</sup>; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 8.03 (1H, s, H-1), 6.65 (1H, s, H-2), 6.43 (1H, br s, -NH<sub>2</sub>); (lit., R<sub>f</sub> 0.33 (EtOAc/hexane (1:2)); mp 99-100 °C)

### 3,4-Dichloropyridin-2(1-*H*)-one (71)



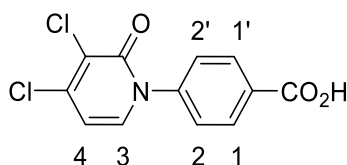
To a stirred solution of 3,4-dichloropyridin-2-amine (170 mg, 1.1 mmol) in conc. HCl (8 mL) at 0°C was added NaNO<sub>2</sub> (150 mg, 2.2 mmol) portionwise. The mixture was stirred at 0°C for 1 hour and then at room temperature for 24 hours. The mixture was slowly poured into a solution of Na<sub>2</sub>CO<sub>3</sub> in water and the pH adjusted to 9, then the aqueous layer was extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 20-80%) gave a colourless oil (88 mg, 16%). R<sub>f</sub> 0.31 (EtOAc/Petrol (2/8)); LCMS (ESI)<sup>+</sup> *m/z* = 164.0 [M + H]<sup>+</sup>; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 7.73 (1H, s, H-1), 6.80 (1H, s, H-2)

#### 4-(3,4-Dichloro-2-oxopyridin-1(2-*H*)-yl)benzonitrile (72)



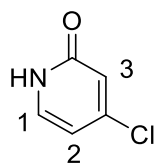
To a stirred solution of 3,4-dichloropyridin-2(1H)-one (38 mg, 0.2 mmol), 4-iodobenzonitrile (58 mg, 0.3 mmol) and  $K_2CO_3$  (48 mg, 0.4 mmol) in anhydrous DMSO (1 mL) was added 8-hydroxyquinoline (8 mg, 0.06 mmol) and copper(I) iodide (13 mg, 0.06 mmol) under  $N_2$  atmosphere. The mixture was stirred at  $130^\circ C$  for 24 hours and then was treated with 20%  $NH_4OH$  aqueous solution and extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried ( $Na_2SO_4$ ), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 50-95%) gave a colourless oil (10 mg, 16%).  $R_f$  0.35 (EtOAc/Petrol (5:95)); ; LCMS (ESI) $^+ m/z = 265.1$  [ $M + H$ ] $^+$ ; UV  $\lambda_{max}$  (EtOH)/nm: 294.2; IR  $\nu_{max}/cm^{-1}$  2300, 2000, 1798, 1610, 1590, 1555, 1407;  $^1H$  NMR (500 MHz, Chloroform-*d*)  $\delta$  8.12 (1H, s, H-4), 7.64 (2H, d,  $J = 8.8$  Hz, H-1, H-1'), 7.16 (2H, d,  $J = 8.8$  Hz, H-2, H-2'), 7.08 (1H, s, H-3);  $^{13}C$  NMR (125 MHz, Chloroform-*d*)  $\delta$  161.0, 156.9, 147.1, 145.0, 134.0 (x2), 126.13, 121.8 (x2), 118.4, 113.7, 108.8; HRMS calc. for  $C_{12}H_6Cl_2N_2O$  [ $M+H$ ] $^+$  264.9930, found 264.9861

#### 4-(3,4-Dichloro-2-oxopyridin-1(2-*H*)-yl)benzoic acid (64)



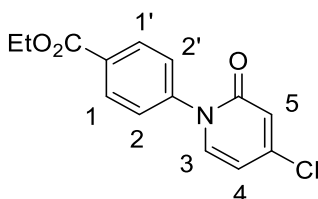
4-(3,4-dichloro-2-oxopyridin-1(2H)-yl)benzonitrile (10 mg, 0.04 mmol) was stirred in a mixture of  $H_2SO_4/H_2O$  (6:4) (1 mL) at room temperature for 16 hours. The mixture was extracted with EtOAc (x3) and the combined organic extracts were washed with brine, dried ( $Na_2SO_4$ ), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4,  $C_{18}$ ,  $CH_3CN$  [0.1% formic acid] in water [0.1% formic acid] eluting 20-80%) gave a yellow oil (0.8 mg, 7%);  $R_f$  0.37 ( $CH_3CN$  [0.1% formic acid] in water [0.1% formic acid], (20:80)); LCMS (ESI) $^+ (m/z) = 266.8$  [ $M + H$ ] $^+$ ;  $^1H$  NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  8.14 (1H, s, H-1), 7.84 (2H, d,  $J = 8.8$  Hz, H-1, H-1'), 7.21 (1H, s, H-3), 7.12 (2H, d,  $J = 8.8$  Hz, H-2, H-2'); Analytical HPLC purity: 100% ( $CH_3CN$  [0.1%  $NH_3$ ])

#### 4-Chloropyridin-2(1-*H*)-one<sup>289</sup> (73)



To a stirred solution of 2,4-dichloropyridine (1g, 6.8 mmoles) in acetic acid (10 mL) was added NaOAc (1.1 g, 13.5 mmoles). The mixture was stirred at 150°C for 24 hours and then was concentrated *in vacuo* affording a yellow solid. The crude product was partitioned between EtOAc and water, the organic layer was separated. Then the aqueous layer was extracted with EtOAc (x4), the combined organic extracts were washed with saturated aqueous solution of NaHCO<sub>3</sub>, brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* to give a white solid (0.9 g, 99.99%). mp: 185-186 °C; LCMS (ESI)<sup>+</sup> *m/z* = 130.0 [M + H]<sup>+</sup>; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 11.86 (1H, s, -NH-), 7.45 (1H, d, *J* = 7.0 Hz, H-1), 6.45 (1H, d, *J* = 2.1 Hz, H-3), 6.27 (1H, d, *J* = 7.0, 2.1 Hz, H-2); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 162.1, 147.0, 137.1, 118.9, 106.8; HRMS calc. for C<sub>5</sub>H<sub>5</sub>ClNO [M+H]<sup>+</sup> 130.0054, found 130.0082; (lit., 184-186 °C; <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 11.82 (br s, 1H, NH), 7.42 (d, 1H, *J* = 7.0 Hz), 6.43 (d, 1H, *J* = 2.1 Hz), 6.25 (dd, 1H, *J* = 7.0 Hz, *J* = 2.1 Hz); <sup>13</sup>C NMR (100 MHz, DMSO-*d*<sub>6</sub>) δ 162.1, 147.0, 137.1, 118.9, 106.7; HRMS (ESI)<sup>+</sup> Calcd for C<sub>5</sub>H<sub>5</sub>ClNO [M + H]<sup>+</sup>: 130.0054, Found: 130.0059)

#### Ethyl-4-(4-chloro-2-oxopyridin-1(2-*H*)-yl)benzoate (74)

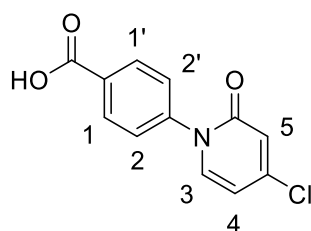


To a stirred solution of 4-chloropyridin-2(1-*H*)-one (0.2 g, 1.6 mmoles), ethyl 4-iodobenzoate (0.5 g, 1.9 mmoles) and *N,N*-dimethylethylenediamine (57 mg, 0.7 mmoles) in anhydrous dioxane (10 mL) was added copper (I) iodide (62 mg, 0.3 mmoles) and K<sub>3</sub>PO<sub>4</sub> (0.7 g, 3.2 mmoles). The mixture was stirred at 110°C for 24 hours and then was filtered through celite and washed with EtOAc. The filtrate was evaporated *in vacuo* and then the residue was taken up with EtOAc, washed with water, brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a brown oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 10%) gave a yellow oil (160 mg, 35%). R<sub>f</sub> 0.31 (EtOAc in Petrol (1/9)); LCMS (ESI)<sup>+</sup> (*m/z*) = 278.0 [M +



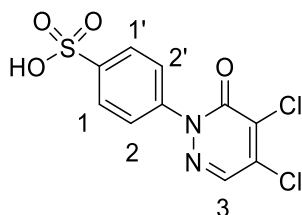
H]<sup>+</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm: 292.6; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 2900, 2250, 2028, 1771, 1578, 1535, 1517, 1455; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  8.09 (1H, d, *J* = 8.6 Hz, H-1), 7.79 (2H, d, *J* = 7.4 Hz, H-5), 7.60 (2H, d, *J* = 8.6 Hz, H-2), 6.70 (1H, d, *J* = 2.3 Hz, H-3), 6.50 (1H, dd, *J* = 7.4, 2.3 Hz, H-4), 4.36 (2H, q, *J* = 7.1 Hz, -CH<sub>2</sub>-), 1.35 (3H, t, *J* = 7.1 Hz, -CH<sub>3</sub>), <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  165.5, 160.4, 146.9, 144.2, 139.9, 130.4 (x2), 130.2, 127.7 (x2), 119.2, 107.8, 61.5, 14.6; HRMS calc. for C<sub>14</sub>H<sub>12</sub>ClNO<sub>3</sub> [M - EtO] 233.0244, found 233.0487

#### 4-(4-Chloro-2-oxopyridin-1(2-*H*)-yl)benzoic acid (65)



To a stirred solution of ethyl 4-(4-chloro-2-oxopyridin-1(2-*H*)-yl)benzoate (0.1 g, 0.4 mmol) in anhydrous THF was added LiOH.H<sub>2</sub>O (60 mg, 0.7 mmol). The mixture was stirred at room temperature for 16 hours and then 0.1 M HCl was added and the pH adjusted to 3. The precipitate formed was filtered through a sintered funnel and dried in a vacuum oven for 16 hours affording a white solid. (48 mg, 53%). *R*<sub>f</sub> 0.58 (EtOAc in Petrol (10/90) + 1 drop of acetic acid); mp 301-303°C; LCMS (ESI)<sup>+</sup> *m/z* = 234.0 [M + H - H<sub>2</sub>O]<sup>+</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm: 314.2; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3370, 1650, 1589, 1511, 1408; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  8.03 (2H, d, *J* = 8.3 Hz, H-1, H-1'), 7.78 (1H, d, *J* = 7.3 Hz, H-5), 7.43 (2H, d, *J* = 8.3 Hz, H-2, H-2'), 6.68 (1H, d, *J* = 2.0 Hz, H-3), 6.47 (1H, d, *J* = 7.3 Hz, H-4), 6.47 (1H, dd, *J* = 7.3, 2.0 Hz, H-4); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  167.9, 160.6, 146.7, 144.2, 140.3, 136.5, 130.3 (x2), 126.7 (x2), 119.1, 107.5; Analytical HPLC purity: 100% (CH<sub>3</sub>CN [0.1% NH<sub>3</sub>]); HRMS calc. for C<sub>12</sub>H<sub>8</sub>ClNO<sub>3</sub> [M+H]<sup>+</sup> 250.0266, found 250.0199

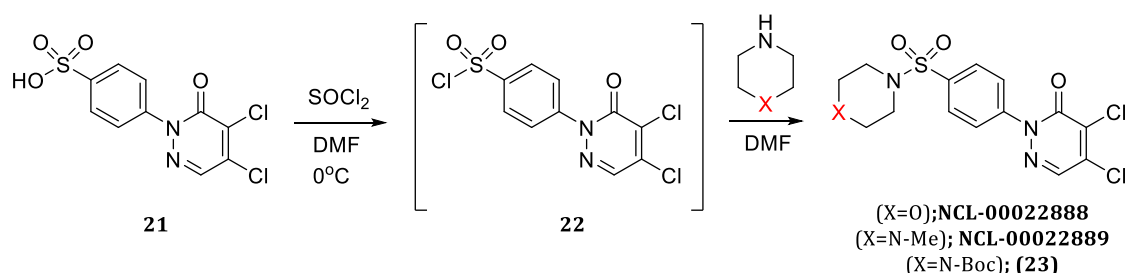
#### 4-(4,5-Dichloro-6-oxopyridazin-1(6-*H*)-yl)benzenesulfonic acid (78)



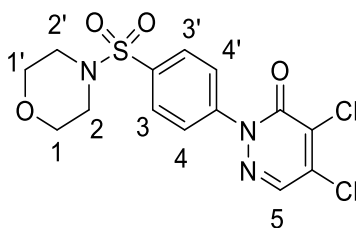
To 4-hydrazinosulfonic acid (2 g, 10.6 mmol) in a mixture of EtOH/H<sub>2</sub>O (1:1) (6 ml) was

added mucochloric acid (1.8 g, 10.6 mmoles) and 0.1 ml of 12 M HCl (catalytic). The mixture was stirred at 65°C for 16 hours and then was filtered through a sintered funnel and washed with cold water affording a pale yellow solid (3,40 g, 99.99%).  $R_f$  0.37 (EtOAc/Petrol/HOAc (9:90:1)); LCMS (ESI)<sup>+</sup>  $m/z$  = 322.0 [M + H]<sup>+</sup>; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  8.35 (1H, s, H-3), 7.73 (2H, d,  $J$  = 8.5 Hz, H-1), 7.51 (2H, d,  $J$  = 8.5 Hz, H-2); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  156.1, 148.9, 141.3, 137.0, 136.4, 134.4, 126.5 (x2), 125.6 (x2)

### General procedure C for (75), (76) and (77)



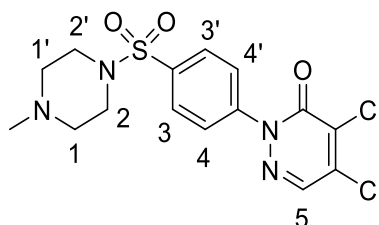
### 4,5-Dichloro-2-(4-(morpholinosulfonyl)phenyl) pyridazin-3(2-*H*)-one (75)



To a solution of 4-(4,5-Dichloro-6-oxopyridazin-1(6-*H*)-yl)benzenesulfonic acid (0.3 g, 0.9 mmoles) in anhydrous DMF (6 mL) was added 150  $\mu$ L of thionyl chloride (0.15 mL, 1.9 mmoles) dropwise at 0°C. After the complete conversion of the acid to the sulphonyl chloride, the mixture was allowed to warm up slowly to room temperature and then morpholine (0.4 mL, 4.8 mmoles) was added dropwise. The mixture was treated with saturated aqueous solution of NaHCO<sub>3</sub> and extracted with EtOAc (x3), the combined organic extracts were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated in vacuo affording a yellow oil that solidified upon standing (160 mg, 66%).  $R_f$  0.38 (EtOAc in Petrol (10/90); mp: 199-200°C; LCMS (ESI)<sup>+</sup>  $m/z$  = 391.0 [M + H]<sup>+</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm: 316.0; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 2961, 2913, 2858, 1672, 1589, 1484, 1446, 1406; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  8.42 (1H, s, H-5), 7.88 – 7.94 (4H, m, H-3, H-3', H-4, H-4'), 3.56 – 3.73 (4H, m, H-2), 2.87 – 2.99 (4H, m, H-1); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  156.1, 145.0, 137.6, 136.6, 134.9, 134.7, 128.9 (x2), 1;

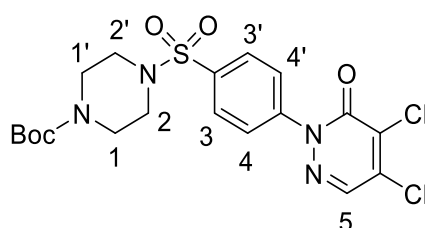
127.1 (x2), 65.8 (x2), 46.4 (x2); Analytical HPLC purity: 100% (CH<sub>3</sub>CN [0.1% NH<sub>3</sub>]); HRMS calc. for C<sub>14</sub>H<sub>13</sub>Cl<sub>2</sub>N<sub>3</sub>O<sub>4</sub>S [M+H]<sup>+</sup> 390.0077, found 390.0120

**4,5-Dichloro-2-(4-((4-methylpiperazin-1-yl)sulfonyl)phenyl)pyridazin-3(2-*H*)-one (76)**



**General procedure C:** Using N-methyl-piperazine (0.31 mL, 2.8 mmol) affording a white solid (17 mg, 5%). *R<sub>f</sub>* 0.41 (EtOAc in Petrol (10/90)); mp: 241-242°C; LCMS (ESI)<sup>+</sup> *m/z* = 404.1 [M + H]<sup>+</sup>; UV λ<sub>max</sub> (EtOH)/nm: 320.0; IR ν<sub>max</sub>/cm<sup>-1</sup> 2938, 2846, 2800, 1676, 1582, 1488, 1453, 1404; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 8.42 (1H, s, H-5); 7.86 – 7.94 (4H, m, H-3, H-3', H-4, H-4'); 2.95 (4H, br s, H-2), 2.33 – 2.40 (4H, m, H-1), 2.15 (3H, s, -CH<sub>3</sub>); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 156.1, 144.9, 137.6, 136.6, 135.3, 134.7, 128.7 (x2), 127.0 (x2), 53.9 (x2), 46.2 (x2), 45.7; Analytical HPLC purity: 100% (CH<sub>3</sub>CN [0.1% NH<sub>3</sub>]); HRMS calc. for C<sub>15</sub>H<sub>16</sub>Cl<sub>2</sub>N<sub>4</sub>O<sub>3</sub>S [M+H]<sup>+</sup> 403.0393, found 403.0311

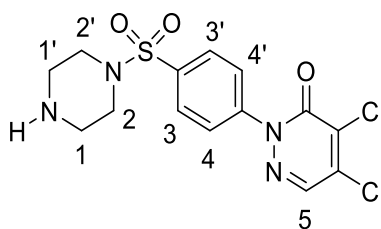
***tert*-Butyl 4-((4-(4,5-dichloro-6-oxopyridazin-1(6-*H*)-yl)phenyl)sulfonyl) piperazine-1-carboxylate (79)**



*tert*-Butyl 4-((4-(4,5-dichloro-6-oxopyridazin-1(6-*H*)-yl)phenyl) sulfonyl)piperazine-1-carboxylate was prepared according to the **general procedure (C)** using N-Boc-piperazine (0.21 g, 1.12 mmol) affording a yellow oil (62 mg, 14%). *R<sub>f</sub>* 0.44 (EtOAc in Petrol (10/90)); LCMS (ESI)<sup>+</sup> *m/z* = 390.1 [M + H - Boc]<sup>+</sup>; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 7.91 (1H, s, H-5), 7.77 – 7.82 (4H, m, H-3, H-3', H-4, H-4'); 3.46 (4H, t, *J* = 5.1 Hz, H-2 (x2), H-2' (x2)); 2.96 (4H, t, *J* = 5.1 Hz, H-1 (x2), H-1' (x2)), 1.35 (9H, s, N-Boc); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 155.9, 154.1, 144.3, 137.0, 136.8, 135.7, 128.4 (x2), 125.6 (x2), 80.5 (x2), 45.9

(x2), 28.3 (x3)

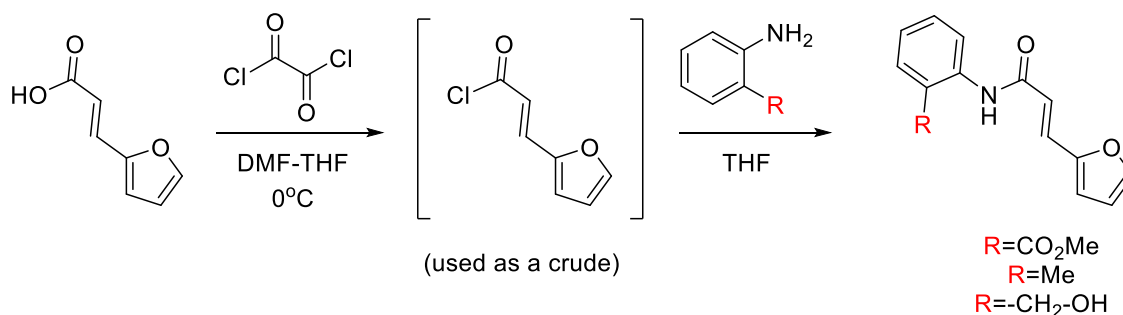
#### 4,5-Dichloro-2-(4-(piperazin-1-ylsulfonyl)phenyl)pyridazin-3(2-*H*)-one (77)



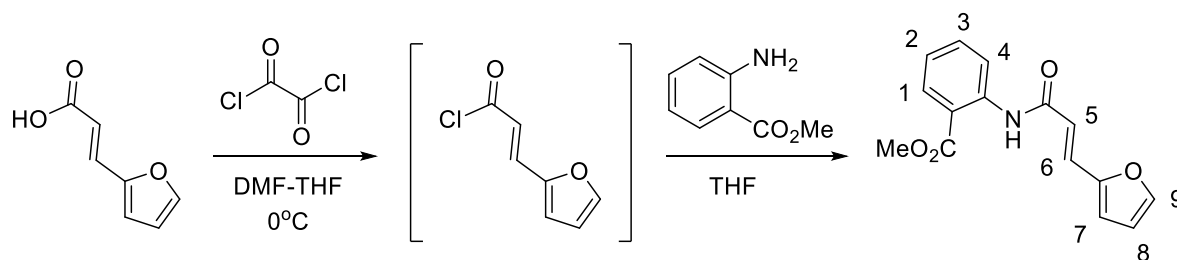
To a stirred solution of 50 mg of *tert*-butyl 4-((4-(4,5-dichloro-6-oxopyridazin-1(6-*H*)-yl)phenyl)sulfonyl) piperazine-1-carboxylate (0.1 mmoles, 1 equiv.) in dioxane (2 mL) and DCM (2 mL) was added 12M HCl (1 mL). The mixture was stirred at room temperature for 24 hours and then was treated with saturated aqueous solution of NaHCO<sub>3</sub>, extracted with EtOAc (x3), the combined organic extracts were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a white solid (17 mg, 35%). *R<sub>f</sub>* 0.71 (EtOAc in Petrol (10/90 + 1 drop of NH<sub>3</sub>)); mp: 236-237°C; LCMS (ESI)<sup>+</sup> *m/z* = 390.1 [M + H]<sup>+</sup>; UV λ<sub>max</sub> (EtOH)/nm: 219.0; IR ν<sub>max</sub>/cm<sup>-1</sup> 3313, 2941, 2903, 2835, 2753, 2317, 1091, 1994, 1918, 1719, 1665, 1586, 1486, 1449, 1404; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 8.41 (1H, s, H-5), 7.96 – 7.84 (4H, m, H-3, H-3', H-4, H-4'); 3.17 (1H, s, -NH-); 2.85 (4H, t, *J* = 4.8 Hz, H-2 (x2), H-2' (x2)); 2.75 (4H, t, *J* = 4.8 Hz, H-1 (x2), H-1' (x2)); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>) δ 156.1, 144.8, 137.5, 136.6, 135.1, 134.7, 128.8 (x2), 127.0 (x2), 47.1 (x2), 45.0 (x2); Analytical HPLC purity: 99.4081% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>14</sub>H<sub>14</sub>Cl<sub>2</sub>N<sub>4</sub>O<sub>3</sub>S [M+H]<sup>+</sup> 389.0236, found 389.0158

### 25.3 Synthesis of (80) and derivatives

#### General procedure D for (80), (81) and (82)

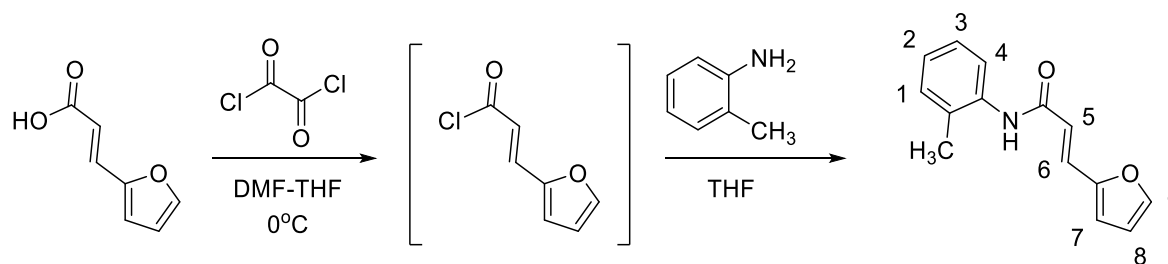


### Methyl-(*E*)-2-(3-(furan-2-yl)acrylamido)benzoate (80)



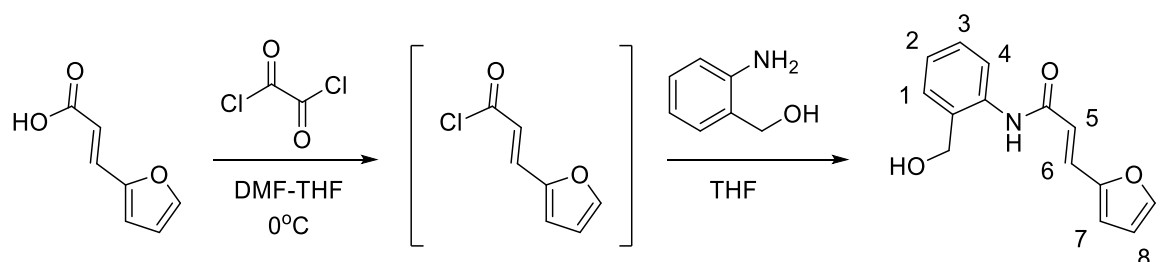
To a solution of (*E*)-3-(furan-2-yl)acrylic acid (0.8 g, 6 mmol) and anhydrous DMF (catalytic) in anhydrous THF (20 mL) was added oxalyl chloride (1.1 mL, 12 mmol) at 0°C. The mixture was stirred at room temperature for 16 hours and then was concentrated to half its volume by rotary evaporation. The concentrate was added dropwise to a stirred solution of methyl 2-aminobenzoate (1.8 g, 6 mmol) and NaHCO<sub>3</sub> (3 g, 18 mmol) in THF/H<sub>2</sub>O (10:1) (60 mL). The mixture was stirred for 4 hours at room temperature and then was evaporated *in vacuo*. The crude product was dissolved in EtOAc and washed with saturated aqueous solution of NaHCO<sub>3</sub>, 1 M aqueous HCl, brine and then dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a pale yellow solid. Chromatography (SP4, silica, EtOAc in Petrol, eluting 10%) gave a white solid (28 mg, 2%). *R<sub>f</sub>* 0.42 (EtOAc in Petrol (10/90)); mp: 224-225°C; UV  $\lambda_{\text{max}}$  (EtOH)/nm: 261.3; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3260, 3131, 2950, 1676, 1588, 1527, 1434; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  10.75 (1H, s, -NH-), 8.29 (1H, dd, *J* = 8.6, 1.2 Hz, H-4), 7.92 (1H, dd, *J* = 8.0, 1.6 Hz, H-1), 7.86 (1H, d, *J* = 1.7 Hz, H-9), 7.63 (1H, ddd, *J* = 8.6, 7.4, 1.6 Hz, H-3), 7.44 (1H, d, *J* = 15.4 Hz, H-6), 7.22 (1H, ddd, *J* = 8.0, 7.4, 1.2 Hz, H-2), 6.94 (1H, d, *J* = 3.3 Hz, H-7), 6.65 (1H, dd, *J* = 3.3, 1.7 Hz, H-8), 6.59 (1H, d, *J* = 15.4 Hz, H-5), 3.86 (3H, s, -CH<sub>3</sub>); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  168.0, 164.0, 151.1, 145.9, 139.8, 134.2, 130.9, 128.7, 123.9, 122.1, 119.5, 119.0, 115.6, 113.2, 52.9; Analytical HPLC purity: 99.5% (CH<sub>3</sub>CN [0.1% NH<sub>3</sub>]); HRMS calc. for C<sub>15</sub>H<sub>13</sub>NO<sub>4</sub> [M + H]<sup>+</sup> 272.0917, found 272.0878

**(E)-3-(Furan-2-yl)-N-(*o*-tolyl)acrylamide (81)**



**General procedure D:** Using *o*-toluidine (820 mg, 5.94 mmol) affording a pale white solid (27 mg, 2%).  $R_f$  0.38 (EtOAc in Petrol (10/90)); mp: 148-149 °C; LCMS (ESI)<sup>+</sup>  $m/z$  = 228.1 [M + H]<sup>+</sup>; UV  $\lambda_{max}$  (EtOH)/nm: 264.5; IR  $\nu_{max}/cm^{-1}$  3258, 3214, 3119, 3010, 2948, 2321; <sup>1</sup>H NMR (500 MHz, Chloroform-*d*)  $\delta$  9.50 (1H, s, -NH-), 7.84 (1H, d,  $J$  = 1.8 Hz, H-9), 7.59 (1H, d,  $J$  = 7.8 Hz, H-4), 7.38 (1H, d,  $J$  = 15.5 Hz, H-6), 7.23 (1H, d,  $J$  = 7.5 Hz, H-1), 7.18 (1H, td,  $J$  = 7.8, 1.5 Hz, H-3), 7.09 (1H, td,  $J$  = 7.5, 1.5 Hz, H-2), 6.85 (1H, d,  $J$  = 3.3 Hz, H-7), 6.78 (1H, d,  $J$  = 15.5 Hz, H-5), 6.63 (1H, dd,  $J$  = 3.3, 1.8 Hz, H-8), 2.24 (3H, s, -CH<sub>3</sub>); <sup>13</sup>C NMR (125 MHz, Chloroform-*d*)  $\delta$  164.0, 151.5, 145.5, 136.8, 131.5, 130.8, 127.6, 126.4, 125.4, 124.9, 119.9, 114.8, 113.0, 18.4; Analytical HPLC purity: 99.5% CH<sub>3</sub>CN [0.1% formic acid]; HRMS calc. for C<sub>14</sub>H<sub>13</sub>NO<sub>2</sub> [M+H]<sup>+</sup> 228.1019, found 228.0979; (lit.,<sup>290</sup> mp 138-140 °C; <sup>13</sup>C NMR (400 MHz, Chloroform-*d*)  $\delta$  17.8, 112.2, 114.4, 117.0, 126.7, 130.1, 130.5, 144.2, 144.5, 151.2, 151.3, 154.0, 166.8)

**(E)-3-(Furan-2-yl)-N-(2-(hydroxymethyl)phenyl)acrylamide (82)**

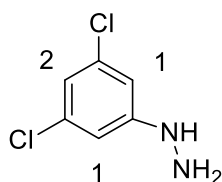


**General procedure D:** Using (2-aminophenyl)methanol (823 mg, 5.96 mmol) affording a pale yellow solid (29 mg, 2%).  $R_f$  0.19 (EtOAc in Petrol (10/90)); mp: 230-231°C; LCMS (ESI)<sup>+</sup>  $m/z$  = 228.1 [M + H]<sup>+</sup>; UV  $\lambda_{max}$  (EtOH)/nm: 268.0; IR  $\nu_{max}/cm^{-1}$  3623, 3038, 2862, 2322, 2105, 1818, 1660, 1620, 1523, 1452; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  9.58 (1H, s, -NH-), 7.89 (1H, d,  $J$  = 1.8 Hz, H-9), 7.71 (1H, d,  $J$  = 7.7 Hz, H-4), 7.50 (1H, d,  $J$  = 7.7 Hz, H-1), 7.44 (1H, d,  $J$  = 15.5 Hz, H-6), 7.31 (1H, t,  $J$  = 7.7, 1.5 Hz, H-3), 7.26 – 7.19 (1H, m, H-2), 6.92 (1H, d,  $J$  = 3.4 Hz, H-7), 6.75 (1H, d,  $J$  = 15.5 Hz, H-5), 6.69 (1H, dd,  $J$  = 3.4, 1.8

Hz, H-8), 5.34 (1H, t,  $J = 5.5$  Hz, -OH), 4.59 (1H, d,  $J = 5.5$  Hz, -CH<sub>2</sub>-); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  164.0, 151.4, 145.6, 135.7, 135.2, 127.79, 127.82, 127.4, 125.2, 124.2, 119.8, 115.0, 113.0, 60.4; Analytical HPLC purity: 98.2% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS (ESI<sup>+</sup>) calc. for C<sub>14</sub>H<sub>13</sub>NO<sub>3</sub> [M+H]<sup>+</sup> 244.0968, found 244.0909

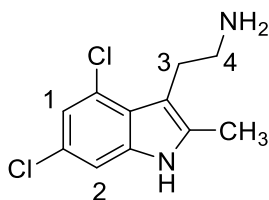
## 25.4 Synthesis of (85) and derivatives

### (3,5-Dichlorophenyl)hydrazine<sup>291</sup> (93)



To a stirred solution of 1,3-dichloroaniline (3 g, 18.5 mmol) in conc. HCl (35 mL) was added NaNO<sub>2</sub> (1.3 g, 18.52 mmol) in water (5 mL) dropwise at -5°C under N<sub>2</sub> atmosphere. The mixture was stirred at 0°C for 30 minutes and then SnCl<sub>2</sub>·(H<sub>2</sub>O)<sub>2</sub> (21.66 g, 96.28 mmol) in conc. HCl (18 mL) was added slowly keeping the reaction mixture below 5°C. The mixture was stirred at 0°C for 1 hour and then was filtered through a pad of celite. The filtrate was evaporated *in vacuo* and the residue was suspended in water, 2M aqueous NaOH was added and the pH adjusted to 9. The aqueous layer was extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, silica, MeOH in DCM, eluting 1-10%) gave a colourless oil (1.6 g, 29%). <sup>1</sup>H NMR (500 MHz, Chloroform-*d*)  $\delta$  7.19 (1H, s, -NH-), 6.68 (1H, t,  $J = 1.8$  Hz, H-2), 6.65 (2H, d,  $J = 1.8$  Hz, -H-1), 5.23 (1H, br s, -NH<sub>2</sub>); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  154.1, 134.9 (x2), 116.4, 109.7 (x2) (lit., <sup>1</sup>H NMR (400 MHz, Chloroform-*d*)  $\delta$  6.75 (t,  $J = 2.0$  Hz, 1H), 6.71 (d,  $J = 2.0$  Hz, 2H), 5.21 (br, 1H)

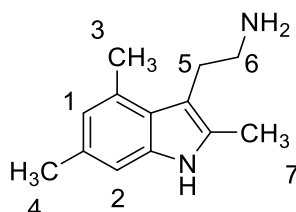
### 2-(4,6-Dichloro-2-methyl-1-*H*-indol-3-yl)ethan-1-amine (85)



To a stirred solution of (3,5-dichlorophenyl)hydrazine (0.1 g, 0.6 mmol) in dry ethanol (2 mL), 5-chloro-2-pentanone (68  $\mu$ L, 0.6 mmol) was added dropwise at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred for 20 minutes and then 0.1 mL of acetic acid

was added. The mixture was stirred at 50°C for 16 hours and then was evaporated *in vacuo*. The residue was triturated in DCM, filtered through a sintered funnel and washed with DCM affording a pale white solid (53 mg, 39%).  $R_f$  0.44 (CH<sub>3</sub>CN [0.1% NH<sub>3</sub>] in water [0.1% NH<sub>3</sub>] (1:1)); UV  $\lambda_{\max}$  (EtOH)/nm: 231.4; IR  $\nu_{\max}$ /cm<sup>-1</sup> 3435, 3284, 2874, 1618, 1552, 1491, 1444; mp: 146-147°C; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  7.25 (1H, d,  $J$  = 1.27 Hz, H-1), 7.02 (1H, d,  $J$  = 1.27 Hz, H-2), 3.28 – 3.21 (2H, m, -CH<sub>2</sub>-3), 3.21 – 3.10 (2H, m, -CH<sub>2</sub>-4), 2.42 (3H, s, -CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  137.2, 135.8, 125.9, 124.1, 123.2, 119.4, 109.3, 105.1, 41.2, 22.6, 9.7; Analytical HPLC purity: 99.1% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>11</sub>H<sub>12</sub>Cl<sub>2</sub>N<sub>2</sub> [M + H]<sup>+</sup> 243.0450, found 243.0371

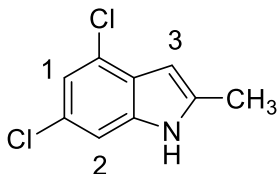
## 2-(2,4,6-Trimethyl-1-*H*-indol-3-yl)ethan-1-amine (87)



To a stirred solution of (3,5-dimethyl)iodobenzene (0.2 g, 0.9 mmoles), *tert*-butyl carbazate (0.14 g, 1 mmol) and caesium carbonate (0.39 g, 1.2 mmoles) in anhydrous DMF (3 mL), CuI (8 mg, 0.04 mmoles) was added at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at 80°C for 2 hours and then was treated with water and extracted with Et<sub>2</sub>O (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a pale brown oil. The crude was dissolved in ethanol (2 mL) and acetic acid (2 mL), and then of 5-chloro-2-pentanone (99  $\mu$ L, 1.7 mmoles) was added dropwise under N<sub>2</sub> atmosphere. The mixture was stirred at 50°C for 16 hours and then was evaporated *in vacuo*. The residue was triturated in DCM, filtered through a sintered funnel and washed with DCM affording an orange oil (14 mg, 8%).  $R_f$  0.29 (CH<sub>3</sub>CN [0.1% NH<sub>3</sub>] in (water [0.1% NH<sub>3</sub>] (1:1)); UV  $\lambda_{\max}$  (EtOH)/nm: 247.0; IR  $\nu_{\max}$ /cm<sup>-1</sup> 3380, 2916, 2855, 1556, 1439; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  6.87 (1H, s, H-1), 6.54 (1H, s, H-2), 2.94 (2H, t,  $J$  = 7.4 Hz, -CH<sub>2</sub>-5), 2.80 (2H, t,  $J$  = 7.4 Hz, -CH<sub>2</sub>-6), 2.59 (3H, s, -CH<sub>3</sub>-7), 2.34 (3H, s, -CH<sub>3</sub>-3,4); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  136.4, 131.1, 129.3, 128.1, 124.5, 121.8, 108.8, 43.8, 28.4, 20.1, 19.0, 10.0; Analytical HPLC purity: 96.2% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>13</sub>H<sub>18</sub>N<sub>2</sub> [M+H]<sup>+</sup> 203.1543, found 203.1468

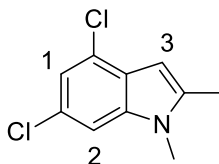


#### 4,6-Dichloro-2-methyl-1-*H*-indole (94)



To a stirred solution of 3,5-dichlorophenylhydrazine (0.125 g, 0.7 mmol) in acetic acid (2 mL) anhydrous acetone (52  $\mu$ L, 0.7 mmol) was added at room temperature under  $N_2$  atmosphere. The mixture was stirred at 120°C for 16 hours and then was evaporated *in vacuo* affording a brown oil. Zinc chloride (0.12 g, 0.84 mmol) was added to the residue and then the mixture was fused at 140°C for 16 hours. The residue was taken up with EtOAc (10 mL), treated with ammonium chloride, washed with brine, dried ( $Na_2SO_4$ ), filtered and evaporated *in vacuo* affording a dark yellow oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 10%) gave a yellow oil (49 mg, 35%).  $R_f$  0.54 (EtOAc in Petrol (5/95)); LCMS (ESI)<sup>+</sup>  $m/z$  = 201.1  $[M + H]^+$ ;  $^1H$  NMR (500 MHz, DMSO- $d_6$ )  $\delta$  7.21 (1H, s, H-1), 6.97 (1H, s, H-2), 6.19 (1H, s, H-3), 2.41 (3H, s, -CH<sub>3</sub>)

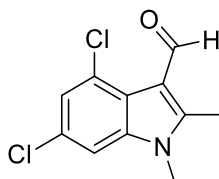
#### 4,6-Dichloro-1,2-dimethyl-1-*H*-indole (95)



To a stirred solution of 4,6-dichloro-2-methyl-1-*H*-indole (49 mg, 0.25 mmol) in anhydrous THF (1 mL) was added NaH (60%) (10 g, 0.27 mmol) at 0°C under  $N_2$  atmosphere. The mixture was stirred at 0°C for 30 minutes and then iodomethane (17  $\mu$ L, 0.27 mmol) was added. The mixture was stirred at 0°C and allowed to warm to room temperature slowly. The mixture was stirred at room temperature for 16 hours and then was treated with H<sub>2</sub>O (3 mL), extracted with DCM (x3), the combined organic extracts were washed with brine, dried ( $Na_2SO_4$ ), filtered and evaporated *in vacuo* affording a pale yellow oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 1%) gave a pale yellow oil (27 mg, 52%);  $R_f$  0.37 (EtOAc in Petrol (1/99)); LCMS (ESI)<sup>+</sup>  $m/z$  = 215.1  $[M + H]^+$ ; UV  $\lambda_{max}$  (EtOH)/nm: 219.3; IR  $\nu_{max}/cm^{-1}$  3023, 2935, 2807, 1665, 1642, 1610;  $^1H$  NMR (500 MHz, Methanol- $d_4$ )  $\delta$  7.25 (1H, dd,  $J$  = 1.6, 1 Hz, H-3), 6.93 (1H, d,  $J$  = 1.6 Hz, H-1), 6.20 (1H, p,  $J$  = 1 Hz, H-2), 3.60 (3H, s, N-CH<sub>3</sub>), 2.36 (3H, d,  $J$  = 1 Hz, -CH<sub>3</sub>);  $^{13}C$  NMR (126 MHz, Methanol- $d_4$ )  $\delta$  119.8, 109.0,

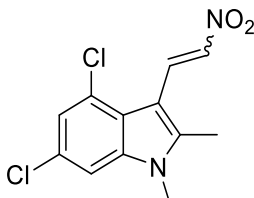
99.0, 30.1, 11.6; HRMS calc. for C<sub>10</sub>H<sub>9</sub>Cl<sub>2</sub>N [M+H]<sup>+</sup> 214.0185, found 214.1468

#### 4,6-Dichloro-1,2-dimethyl-1-*H*-indole-3-carbaldehyde (96)



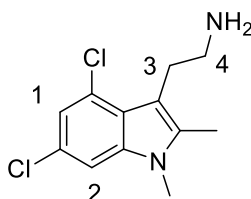
POCl<sub>3</sub> was added dropwise to a stirred solution of anhydrous DMF (1 mL) at 0°C under N<sub>2</sub> atmosphere. The mixture was stirred at 0°C for 20 minutes and then 4,6-dichloro-1,2-dimethyl-1-*H*-indole (29 mg, 0.14 mmol) in anhydrous DMF (1 mL) was added dropwise at 0°C. The mixture stirred for 1 hour at room temperature and then was treated with aqueous 2M NaOH and the pH adjusted to 12. The aqueous layer was extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a pale pink oil (33mg, 99.99%) which was used in the next step without further purification due to stability issues.

#### 4,6-Dichloro-1,2-dimethyl-3-(2-nitrovinyl)-1-*H*-indole (97)



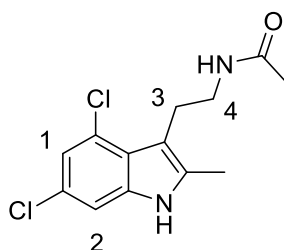
To a stirred solution of 4,6-dichloro-1,2-dimethyl-1H-indole-3-carbaldehyde (33 mg, 0.14 mmol) in CH<sub>3</sub>NO<sub>2</sub> (3 mL) ammonium acetate (37 mg, 0.48 mmol) was added at room temperature. The mixture was stirred at 80°C for 16 hours and then was evaporated *in vacuo*. The residue was taken up with EtOAc, washed with water, brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording an orange oil (39 mg, 99.99%) which was used in the next step without further purification.

## 2-(4,6-Dichloro-1,2-dimethyl-1-*H*-indol-3-yl)ethan-1-amine (89)



To a stirred solution of lithium aluminium hydride (20 mg, 0.05 mmoles) in anhydrous THF (1.5 mL) was added 4,6-dichloro-1,2-dimethyl-3-(2-nitrovinyl)1-*H*-indole (39 mg, 0.14 mmoles) in anhydrous THF (1 mL) dropwise at 0°C under N<sub>2</sub> atmosphere. The mixture was stirred at room temperature for 16 hours and then was quenched with water, extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a pale yellow oil. Chromatography (SP4, C<sub>18</sub>, CH<sub>3</sub>CN [0.1% NH<sub>3</sub>] in water [0.1% NH<sub>3</sub>], eluting 33%) gave a colourless oil (1 mg, 3%). *R*<sub>f</sub> 0.41 (CH<sub>3</sub>CN [0.1% NH<sub>3</sub>] in water [0.1% NH<sub>3</sub>] (1:1)); LCMS (ESI)<sup>+</sup> *m/z* = 258.1 [M + H]<sup>+</sup>; UV λ<sub>max</sub> (EtOH)/nm: 227.6; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 7.33 (1H, d, *J* = 1.7 Hz, H-1), 7.00 (1H, d, *J* = 1.7 Hz, H-2), 3.68 (3H, s, N-CH<sub>3</sub>), 3.13 (2H, t, *J* = 7.5 Hz, -CH<sub>2</sub>-4), 2.93 (2H, t, *J* = 7.5 Hz, -CH<sub>2</sub>-3) 2.41 (3H, s, -CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>) δ 139.7, 138.3, 127.1, 126.1, 124.1, 120.8, 109.1, 108.9, 44.3, 30.2, 28.0, 10.1; Analytical HPLC purity: 91.7% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>12</sub>H<sub>14</sub>Cl<sub>2</sub>N<sub>2</sub> [M+H]<sup>+</sup> 257.0607, found 257.1468

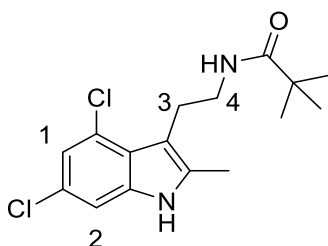
## *N*-(2-(4,6-Dichloro-2-methyl-1-*H*-indol-3-yl)ethyl)acetamide (90)



To a stirred solution of 2-(4,6-dichloro-2-methyl-1H-indol-3-yl)ethan-1-amine (93 mg, 0.4 mmoles) in acetonitrile (2 mL) and anhydrous DMF (1 mL), triethylamine (0.265 mL, 1.9 mmoles) was added dropwise at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at room temperature for 20 minutes and then acetic anhydride (41 μL, 0.43 mmoles) was added. The mixture was stirred at room temperature for 16 hours and then was treated with saturated aqueous solution of NaHCO<sub>3</sub> and extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo*

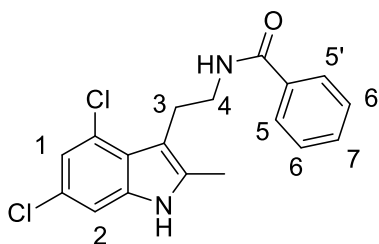
affording a yellow oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 45-100%) gave a white solid (97 mg, 89%).  $R_f$  0.21 (EtOAc in Petrol (2:1)); UV  $\lambda_{\max}$  (EtOH)/nm: 287.0; IR  $\nu_{\max}/\text{cm}^{-1}$  3281, 2933, 2400, 1608, 1437; mp: 155-156°C;  $^1\text{H}$  NMR (500 MHz, Methanol- $d_4$ )  $\delta$  8.09 (1H, s, -NH-), 7.17 (1H, d,  $J$  = 1.7 Hz, H-1), 6.92 (1H, d,  $J$  = 1.7 Hz, H-2), 3.38 (2H, t,  $J$  = 6.3 Hz, -CH<sub>2</sub>-4), 3.35 (s, 3H), 3.05 (2H, t,  $J$  = 6.3 Hz, CH<sub>2</sub>-3), 2.34 (3H, s, -CH<sub>3</sub>), 1.91 (3H, s, -C(=O)-CH<sub>3</sub>);  $^{13}\text{C}$  NMR (126 MHz, Methanol- $d_4$ )  $\delta$  173.2, 138.5, 136.6, 126.7, 126.0, 125.0, 120.5, 110.3, 109.3, 42.6, 25.5, 22.6, 11.2; Analytical HPLC purity: 99.1% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>13</sub>H<sub>14</sub>Cl<sub>2</sub>N<sub>2</sub>O [M+H]<sup>+</sup> 285.0556, found 285.0477

***N*-(2-(4,6-dichloro-2-methyl-1-*H*-indol-3-yl)ethyl)pivalamide (91)**



To a stirred solution of 2-(4,6-dichloro-2-methyl-1-*H*-indol-3-yl)ethan-1-amine (0.15 g, 0.62 mmoles) in anhydrous acetonitrile (3 mL) and anhydrous DMF (1 mL), triethylamine (427  $\mu\text{L}$ , 3.1 mmoles) was added dropwise at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at room temperature for 20 minutes and then pivalic anhydride (140  $\mu\text{L}$ , 0.69 mmoles) was added. The mixture was stirred at room temperature for 16 hours and then was treated with saturated aqueous solution of Na<sub>2</sub>CO<sub>3</sub> and extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 25-75%) gave a pale white solid (97 mg, 48%).  $R_f$  0.51 (EtOAc in Petrol (2:1)); UV  $\lambda_{\max}$  (EtOH)/nm: 287.8; IR  $\nu_{\max}/\text{cm}^{-1}$  3447, 3206, 2934, 1639, 1523, 1433; mp: 184-185°C;  $^1\text{H}$  NMR (500 MHz, Chloroform- $d$ )  $\delta$  9.30 (1H, s, -NH-C(=O)), 7.14 (1H, d,  $J$  = 1.7 Hz, H-1), 7.00 (1H, d,  $J$  = 1.7 Hz, H-2), 3.52 (2H, q,  $J$  = 6.9 Hz, -CH<sub>2</sub>-4), 3.09 (2H, t,  $J$  = 6.9 Hz, -CH<sub>2</sub>-3), 2.28 (3H, s, -CH<sub>3</sub>), 1.13 (9H, s, -(CH<sub>3</sub>)<sub>3</sub>);  $^{13}\text{C}$  NMR (126 MHz, Chloroform- $d$ )  $\delta$  179.0, 136.9, 134.9, 126.1, 124.9, 123.9, 120.4, 109.6, 108.5, 41.4, 38.7, 27.6 (x3), 24.3, 11.6; Analytical HPLC purity: 100% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>13</sub>H<sub>14</sub>Cl<sub>2</sub>N<sub>2</sub>O [M+H]<sup>+</sup> 327.1026, found 327.0944

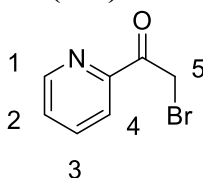
## N-(2-(4,6-dichloro-2-methyl-1H-indol-3-yl)ethyl)benzamide (92)



To a stirred solution of 2-(4,6-dichloro-2-methyl-1H-indol-3-yl)ethan-1-amine (100 mg, 0.41 mmol) in CH<sub>3</sub>CN/DMF (3:1) (4 mL), triethylamine (0.29 mL, 2 mmol) was added dropwise at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at room temperature for 20 minutes and then benzoyl chloride (54  $\mu$ L, 0.46 mmol) was added. The mixture was stirred at room temperature for 16 hours and then was treated with saturated aqueous solution of NaHCO<sub>3</sub>, extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 15-50%) gave a white solid (10 mg, 7%). *R*<sub>f</sub> 0.23 (EtOAc in Petrol (1:3)); UV  $\lambda_{\text{max}}$  (EtOH)/nm: 235.0; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3902, 3869, 3854, 3838, 3748, 3736, 3674, 3649, 3629, 3589, 3566, 1650, 1635, 1541, 1521, 1455, 1420; mp: 283-284°C; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  7.80 – 7.75 (2H, m, H-5, H-5'), 7.56 – 7.49 (1H, m, H-7), 7.47 – 7.41 (2H, m, H-6, H-6'), 7.19 (1H, d, *J* = 1.7 Hz, H-1), 6.96 (1H, d, *J* = 1.7 Hz, H-2), 3.64 (2H, t, *J* = 7.1 Hz, -CH<sub>2</sub>-4), 3.22 (2H, t, *J* = 7.1 Hz, -CH<sub>2</sub>-3), 2.33 (3H, s, -CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  170.3, 138.5, 136.7, 135.9, 132.5, 129.5 (x2), 128.2 (x2), 126.6, 126.0, 125.1, 120.5, 110.3, 109.4, 43.1, 25.4, 11.3; Analytical HPLC purity: 97.9% (CH<sub>3</sub>CN [0.1% NH<sub>3</sub>]); HRMS calc. for C<sub>18</sub>H<sub>16</sub>Cl<sub>2</sub>N<sub>2</sub>O [M+H]<sup>+</sup> 347.0713, found 347.0144

## 25.5 Synthesis of (86) and derivatives

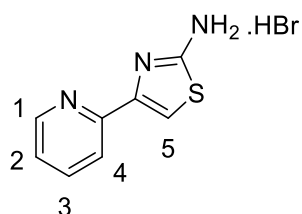
### 2-Bromo-1-(pyridin-2-yl)ethan-1-one<sup>292</sup> (107)



To a stirred solution of 2-acetyl-pyridine (250 mg, 2 mmol) in 45% hydrobromic acid in acetic acid (747  $\mu$ L, 6.2 mmol) bromine (181 mg, 2.27 mmol) was added at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at room temperature for 16 hours

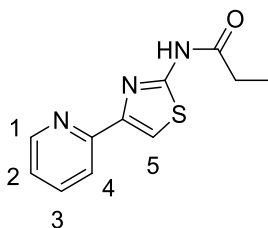
and then treated with Et<sub>2</sub>O (8 mL). The precipitate formed was filtered through a pad of celite, washed with Et<sub>2</sub>O and the filtrate evaporated *in vacuo* affording a pale yellow solid (702 mg, 85%) which was used in the next step without further purification.; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 8.83 (1H, dd, *J* = 5.8, 1.4 Hz, H-1), 8.75 (1H, td, *J* = 8.1, 1.4 Hz, H-3), 8.31 (d, *J* = 8.1 Hz, H-4), 8.22 – 8.16 (m, H-2), 3.95 (1H, d, *J* = 11.5 Hz, -CH<sub>2</sub>-), 3.85 (1H, *J* = 11.5 Hz, -CH<sub>2</sub>-); <sup>13</sup>C NMR (126 MHz, DMSO-*d*<sub>6</sub>) δ 200.0, 149.5, 138.5, 128.7, 121.8, 65.6, 49.1; (lit., <sup>1</sup>H NMR (Methanol-*d*<sub>4</sub>) δ 8.81 (1H, d, *J* = 5.4 Hz), 8.73 (1H, t, *J* = 8.1 Hz), 8.27 (1H, d, *J* = 6.1 Hz), 8.14 (1H, t, *J* = 6.6 Hz), 3.92 (1H, d, *J* = 11.4 Hz), 3.83 (1H, d, *J* = 11.4 Hz))

#### 4-(Pyridin-2-yl)thiazol-2-amine<sup>293</sup> (108)



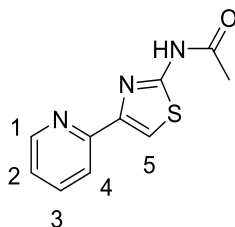
To a stirred solution of 2-bromo-1-(pyridine-2-yl)ethan-1-one (100 mg, 0.5 mmoles) in dry EtOH (2 mL) was added thiourea (38 mg, 0.55 mmoles) at room temperature. The mixture was stirred at room temperature for 5 hours and then was cooled down to 0°C. The precipitate formed was filtered through a sintered funnel and then dissolved in 2M aqueous NaOH (4 mL) extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a pale yellow solid (105 mg, 81%) which was used in the next step without further purification. <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 8.48 – 8.59 (1H, m, H-1), 7.78 – 7.90 (2H, m, H-4, H-3), 7.24 – 7.37 (2H, m, H-5, H-2), 7.18 (2H, br s, -NH<sub>2</sub>); (lit., <sup>1</sup>H NMR (600 MHz, Chloroform-*d*) δ 8.58 (1H, d, *J* = 4.4 Hz), 7.88 (1H, d, *J* = 7.9 Hz), 7.71 (1H, td, *J* = 7.6, 1.5 Hz), 7.31 (1H, s), 7.18 (1H, ddd, *J* = 7.3, 4.7, 1.2 Hz), 5.05 (2H br s))

#### N-(4-(Pyridin-2-yl)thiazol-2-yl)propionamide (86)



To a stirred solution of 4-(pyridin-2-yl)thiazol-2-amine (50 mg, 0.2 mmol) and anhydrous pyridine (101  $\mu$ L, 1.25 mmol) in dioxane (2 mL) was added propionyl chloride (17  $\mu$ L, 0.2 mmol) dropwise at 0°C under N<sub>2</sub> atmosphere. The mixture was slowly warmed up to room temperature for 30 minutes and stirred for 5 hours. The mixture was treated with saturated aqueous solution of NaHCO<sub>3</sub> (3 mL), extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a pale yellow oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 30-70%) gave a white solid (30 mg, 67%). *R*<sub>f</sub> 0.62 (EtOAc in Petrol (1:1)); UV  $\lambda_{\text{max}}$  (EtOH)/nm: 234.0, 269; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 2982, 2229, 1670, 1501; mp: 226-227°C; <sup>1</sup>H NMR (500 MHz, Chloroform-*d*)  $\delta$  9.51 (1H, br s, -NH-), 8.63 (1H, d, *J* = 4.6, 1.7 Hz H-1), 7.89 (1H, d, *J* = 7.8, 1.2 Hz, H-4), 7.74 (td, *J* = 7.8, 1.7 Hz, H-3), 7.68 (1H, s, H-5), 7.22 (1H, ddd, *J* = 7.8, 4.6, 1.2 Hz, H-2), 2.42 (2H, q, *J* = 7.6 Hz, -CH<sub>2</sub>-), 1.22 (3H, t, *J* = 7.6 Hz, -CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, Chloroform-*d*)  $\delta$  171.5, 158.2, 152.3, 149.7, 149.4, 136.9, 122.7, 120.5, 112.0, 29.4, 9.0; Analytical HPLC purity: 100% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>11</sub>H<sub>11</sub>N<sub>3</sub>OS [M+H]<sup>+</sup> 234.0696, found 234.0631

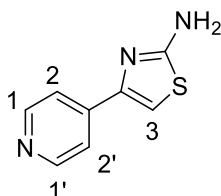
#### ***N*-(4-(pyridin-2-yl)thiazol-2-yl)acetamide (104)**



To a stirred solution of 4-(pyridin-2-yl)thiazol-2-amine (137 mg, 0.62 mmol) and anhydrous pyridine (252  $\mu$ L, 3.1 mmol) in dioxane (2 mL) was added acetic anhydride (66  $\mu$ L, 0.7 mmol) dropwise at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at room temperature for 5 hours and then was treated with saturated aqueous solution of Na<sub>2</sub>CO<sub>3</sub> (5 mL), extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a pale yellow oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 35%) to give a white solid (30 mg, 67%). *R*<sub>f</sub> 0.55 (EtOAc in Petrol (1:1)); UV  $\lambda_{\text{max}}$  (EtOH)/nm: 268.3, 234.6; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3166, 3065, 2993, 1639, 1552, 1424; mp: 249-250°C; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  12.29 (1H, s, -NH-), 8.60 (1H, ddd, *J* = 4.8, 1.8, 1.0 Hz, H-1), 7.93 (dt, *J* = 7.7, 1.0 Hz, H-4), 7.89 (td, *J* = 7.7, 1.8 Hz, H-3), 7.80 (s, H-5), 7.33 (ddd, *J* = 7.7, 4.8, 1.0 Hz, H-2), 2.18 (3H, s, -CH<sub>3</sub>);

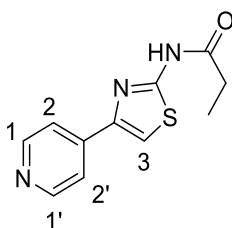
$^{13}\text{C}$  NMR (126 MHz,  $\text{DMSO-}d_6$ )  $\delta$  169.2, 158.7, 152.5, 150.0, 149.4, 137.7, 123.3, 120.4, 111.9, 23.0; Analytical HPLC purity: 100% ( $\text{CH}_3\text{CN}$  [0.1% formic acid]); HRMS calc. for  $\text{C}_{10}\text{H}_9\text{N}_3\text{OS}$   $[\text{M}+\text{H}]^+$  220.0539, found 220.0461

#### 4-(Pyridin-4-yl)thiazol-2-amine<sup>294</sup> (111)



To a stirred solution of 2-bromo-1-(pyridin-4-yl)ethan-1-one (0.5 g, 0.85 mmol) in EtOH (10 mL) was added thiourea (71 mg, 0.92 mmol) portionwise at room temperature under  $\text{N}_2$  atmosphere. The mixture was stirred at room temperature for 6 hours and then the precipitate formed was filtered through a sintered funnel, dissolved in 2M aqueous NaOH (5 mL), extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried ( $\text{Na}_2\text{SO}_4$ ), filtered and evaporated *in vacuo* to afford a white solid (400 mg, 60%) which was used in the next step without further purification. m.p. 245-246  $^\circ\text{C}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{DMSO-}d_6$ )  $\delta$  8.54 (2H, d,  $J = 6.1$  Hz, H-1, H-1'), 7.72 (d,  $J = 6.1$  Hz, H-2, H-2'), 7.38 – 7.43 (1H, m, H-3), 7.21 (2H, br s,  $-\text{NH}_2$ );  $^{13}\text{C}$  NMR (126 MHz,  $\text{DMSO-}d_6$ )  $\delta$  169.0, 150.5 (x2), 147.9, 141.9, 120.3 (x2), 106.6; (lit., mp 278-282  $^\circ\text{C}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{Methanol-}d_4$ )  $\delta$  8.49 (2H, dd,  $J = 1.6, 4.7$  Hz), 7.79 (2H, dd,  $J = 1.6, 4.7$  Hz), 7.25 (1H, s))

#### N-(4-(Pyridin-4-yl)thiazol-2-yl)propionamide (112)

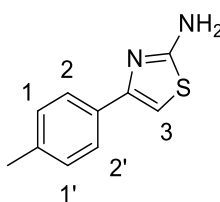


To a stirred solution of 4-(pyridin-4-yl)thiazol-2-amine (75 mg, 0.42 mmol) and anhydrous pyridine (171  $\mu\text{L}$ , 2.1 mmol) in anhydrous DCM (3 mL) was added propionyl chloride (42  $\mu\text{L}$ , 0.47 mmol) dropwise at 0 $^\circ\text{C}$  under  $\text{N}_2$  atmosphere. The mixture was allowed to warm to room temperature slowly and stirred at room temperature for 5 hours. The mixture was treated with saturated aqueous solution of  $\text{Na}_2\text{CO}_3$  (5 mL) and extracted with EtOAc (x3), the



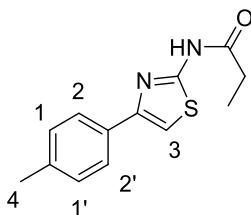
combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* to afford a white solid (78 mg, 80%). *R<sub>f</sub>* 0.59 (EtOAc in Petrol (1:1)); UV  $\lambda_{\text{max}}$  (EtOH)/nm: 280.6, 229.2; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3166, 3065, 2993, 1639, 1552, 1424; mp: 251-252°C; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  12.32 (1H, s, -NH-), 8.62 (2H, d, *J* = 6.1 Hz, H-1), 7.98 (1H, d, *J* = 0.9 Hz, H-3), 7.84 (2H, d, *J* = 6.1 Hz, H-2), 2.47 (2H, q, *J* = 7.5 Hz, -CH<sub>2</sub>-), 1.11 (3H, t, *J* = 7.5 Hz, -CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  173.0, 159.0, 150.7 (x2), 146.7, 141.5, 120.4 (x2), 112.6, 28.7, 9.6; Analytical HPLC purity: 100% (CH<sub>3</sub>CN [0.1% formic acid]); HRMS calc. for C<sub>11</sub>H<sub>11</sub>N<sub>3</sub>OS [M+H]<sup>+</sup> 234.0696, found 234.0618

#### 4-(*p*-Tolyl)thiazol-2-amine<sup>295</sup> (113)



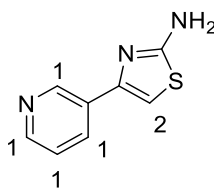
To a stirred solution of 2-bromo-1-(*p*-tolyl)ethan-1-one (0.5 g, 2.16 mmol) in dry EtOH (10 mL) was added thiourea (181 mg, 2.32 mmol) portionwise at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at room temperature under N<sub>2</sub> atmosphere for 6 hours, then the precipitate formed was filtered through a pad of celite and washed with EtOH. The filtrate was evaporated *in vacuo*. The residue was dissolved in 2M aqueous NaOH (5 mL), extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a white solid (210 mg, 44%) which was used in the next step without further purification. UV  $\lambda_{\text{max}}$  (EtOH)/nm: 212.2; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3289, 3174, 3136, 3083, 2999, 2110; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  8.71 (2H, br s, -NH<sub>2</sub>), 7.63 (2H, d, *J* = 8.0 Hz, H-1, H-1'), 7.30 (d, *J* = 8.0 Hz, H-2, H-2'), 7.17 (1H, s, H-3), 2.35 (3H, s, -CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  170.5, 139.4 (x2), 130.0, 126.2 (x2), 102.2, 21.3 (2 missing carbon signals); (lit., <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 400 MHz)  $\delta$  8.87 (2H, br s, NH<sub>2</sub>), 7.48-7.41 (4H, m, Ar-H), 6.81 (1H, s, -CH), 2.31 (3H, s, Me), <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>, 75 MHz):  $\delta$  169.8, 139.2, 138.2, 129.0, 124.8, 124.3, 99.1, 20.49)

#### ***N*-(4-(*p*-Tolyl)thiazol-2-yl)propionamide (114)**



To a stirred solution of 2-bromo-1-(*p*-tolyl)ethan-1-one (0.2 g, 1.1 mmol) and anhydrous pyridine (170  $\mu$ L, 2.1 mmol) in anhydrous DCM (5 mL) was added propionyl chloride (95  $\mu$ L, 1.1 mmol) dropwise at room temperature under  $N_2$  atmosphere. The mixture was stirred at room temperature for 16 hours, then was treated with saturated aqueous solution of  $Na_2CO_3$  (5 mL) and extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried ( $Na_2SO_4$ ), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, silica EtOAc in Petrol, eluting 40%) gave a white solid (30 mg, 80%).  $R_f$  0.62 (EtOAc in Petrol (1:1)); UV  $\lambda_{max}$  (EtOH)/nm: 268.8, 235.8, 210.6; IR  $\nu_{max}/cm^{-1}$  3130, 3023, 2972, 2913, 2872, 1685, 1534; mp: 241-242 $^{\circ}C$ ;  $^1H$  NMR (500 MHz, Methanol- $d_4$ )  $\delta$  7.79 (2H, d,  $J = 8.0$  Hz, H-2), 7.29 (1H, s, H-3), 7.22 (2H, d,  $J = 8.0$  Hz, H-1), 2.52 (2H, q,  $J = 7.6$  Hz,  $-CH_2-$ ), 2.37 (3H, s,  $-CH_3$ -4), 1.25 (3H, t,  $J = 7.6$  Hz,  $-CH_3$ );  $^{13}C$  NMR (126 MHz, Methanol- $d_4$ )  $\delta$  173.1, 158.0, 149.8, 137.3, 131.9, 128.8 (x2), 125.6 (x2), 106.2, 28.4, 19.8, 8.2; Analytical HPLC purity: 100% ( $CH_3CN$  [0.1% formic acid]; HRMS calc. for  $C_{13}H_{14}N_2OS$   $[M+H]^+$  247.0900, found 247.1830

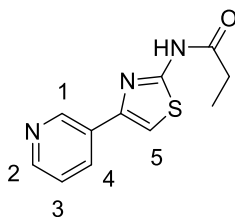
#### **4-(Pyridin-3-yl)thiazol-2-amine<sup>296</sup> (108)**



To a stirred solution of 2-bromo-1-(pyridin-3-yl)ethan-1-one (0.5 g, 0.85 mmol) in dry EtOH (10 mL) was added thiourea (71 mg, 0.92 mmol) portionwise at room temperature under  $N_2$  atmosphere. The mixture was stirred at room temperature for 6 hours. The precipitate formed was filtered through a pad of celite and washed with EtOH. The filtrate was evaporated *in vacuo* and the residue dissolved in 2M NaOH (5 mL), extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried ( $Na_2SO_4$ ), filtered

and evaporated *in vacuo* affording a white solid (109 mg, 16%) which was used in the next step without further purification.  $^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ )  $\delta$  7.45-9.05 (4H, m, Ar-H-1), 7.19 (1H, s, H-2), 3.55 (2H, s, -NH $_2$ );  $^{13}\text{C}$  NMR (126 MHz, DMSO- $d_6$ )  $\delta$  173.0, 149.2, 148.5, 135.6, 124.0, 105.9; (lit.,  $^1\text{H}$  NMR (DMSO- $d_6$ )  $\delta$  7.37-9.01 (m, 4H, Ar-H), 7.19 (s, 1H, CH of thiazole ring), 3.52 (2H, s, -NH $_2$ );  $^{13}\text{C}$  NMR (DMSO- $d_6$ )  $\delta$  169.05, 148.40, 147.28, 132.97, 130.81, 123.99, 103.49

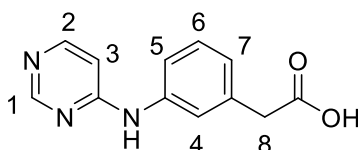
### ***N*-(4-(pyridin-3-yl)thiazol-2-yl)propionamide (109)**



To a stirred solution of 4-(pyridin-3-yl)thiazol-2-amine (0.1 g, 0.56 mmoles) and anhydrous pyridine (228  $\mu\text{L}$ , 2.8 mmoles) in anhydrous DCM (3 mL) was added propionyl chloride (56  $\mu\text{L}$ , 0.6 mmoles) dropwise at room temperature under N $_2$  atmosphere. The mixture was stirred at room temperature for 16 hours, then was treated with saturated aqueous solution of NaHCO $_3$  (5 mL) and extracted with EtOAc (x3), the combined organic extracts were washed with brine, dried (Na $_2$ SO $_4$ ), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 35%) gave a white solid (15 mg, 11%).  $R_f$  0.66 (EtOAc in Petrol (1:1)); UV  $\lambda_{\text{max}}$  (EtOH)/nm: 231.2, 269.2; IR  $\nu_{\text{max}}$ /cm $^{-1}$  3092, 3054, 2979, 2726, 2106, 1681, 1576, 1505, 1461, 1408; mp: 271-272°C;  $^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ )  $\delta$  12.28 (1H, s, -NH-), 9.11 (1H, d,  $J$  = 2.7 Hz, H-1), 8.53 (dd,  $J$  = 4.8, 1.6 Hz, H-2), 8.23 (dt,  $J$  = 8.0, 1.8 Hz, H-4), 7.79 (1H, s, H-5), 7.47 (1H, ddd,  $J$  = 8.0, 4.8, 0.7 Hz, H-3), 2.47 (2H, q,  $J$  = 7.5 Hz, -CH $_2$ -), 1.11 (3H, t,  $J$  = 7.5 Hz, -CH $_3$ );  $^{13}\text{C}$  NMR (126 MHz, DMSO- $d_6$ )  $\delta$  172.9, 159.0, 149.1, 147.4, 146.4, 133.3, 130.5, 124.3, 109.9, 28.7, 9.6; Analytical HPLC purity: 98.9% (CH $_3$ CN [0.1% formic acid]); HRMS calc. for C $_{11}$ H $_{11}$ N $_3$ OS [M+H] $^+$  234.0696, found 234.6240

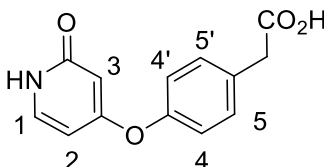
## **25.6 Synthesis of FragLites derivatives**

## 2-(3-(pyrimidin-4-ylamino)phenyl)acetic acid (115)



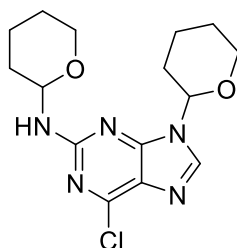
To a solution of 4-chloropyrimidine (100 mg, 0.9 mmol) in anhydrous DMF (4 mL) was added 2-(3-aminophenyl) acetic acid (100 mg, 0.66 mmol). The mixture was stirred at 80°C for 3 hours, and then was evaporated *in vacuo*. Chromatography (SP4, silica, MeOH [0.1% AcOH] in DCM, eluting 10%) gave a beige solid (15 mg, 10%).  $R_f$  0.47 (DCM:MeOH:AcOH 9:1:0.05), IR  $\nu_{\max}/\text{cm}^{-1}$  3045, 2827, 1649, 1580, 1516, 1474; mp 186-187 °C;  $^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ )  $\delta$  12.40 (1H, br s, -CO<sub>2</sub>H), 9.89 (1H, s, H-1), 8.65 (1H, s, H-2), 8.28 (1H, s, H-4), 7.63 (1H, d,  $J$  = 7.8 Hz, H-5), 7.53 (1H, s, -NH-), 7.28 (1H, t,  $J$  = 7.8 Hz, H-6), 6.95 (1H, d,  $J$  = 7.8 Hz, H-7), 6.85 (1H, d,  $J$  = 5.7 Hz, H-3), 3.56 (2H, s, -CH<sub>2</sub>-8);  $^{13}\text{C}$  NMR (500 MHz, DMSO- $d_6$ )  $\delta$  159.9, 158.0, 155.1, 139.6, 128.6, 123.5, 120.6, 118.2, 107.2, 41.0 (1 missing quaternary carbon -CO<sub>2</sub>H); HRMS calc. for C<sub>12</sub>H<sub>11</sub>N<sub>3</sub>O<sub>2</sub> [M + H]<sup>+</sup> 230.0924; found 230.0925

## 2-(4-((2-oxo-1,2-dihydropyridin-4-yl)oxy)phenyl)acetic acid (117)



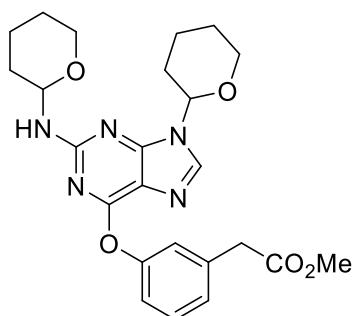
A mixture of 2-(4-((2-methoxypyridine-4-yl)oxy)phenyl)acetate (57 mg, 0.2 mmol) and pyridine hydrochloride (146 mg, 1.2 mmol) were stirred at 120°C for 24 hours. The mixture was allowed to cool to room temperature, and then was evaporated *in vacuo*. Chromatography (SP4, C<sub>18</sub>, CH<sub>3</sub>CN in H<sub>2</sub>O, eluting 10-40%) gave a beige solid (19 mg, 37%).  $R_f$  0.61 (MeOH/DCM (1:9)), IR  $\nu_{\max}/\text{cm}^{-1}$  3443, 1613, 1216, 1182; mp 260-261 °C;  $^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ )  $\delta$  11.35 (1H, s, -CO<sub>2</sub>H), 7.36 (1H, d,  $J$  = 7.3 Hz, H-3), 7.34 (2H, d,  $J$  = 8.4 Hz, H-5, H-5'), 7.07 (2H, d,  $J$  = 8.4 Hz, H-4, H-4'), 5.97 (1H, dd,  $J$  = 7.3, 2.5 Hz, H-2), 5.31 (1H, d,  $J$  = 2.5 Hz, H-1), 3.53 (2H, s, -CH<sub>2</sub>-);  $^{13}\text{C}$  NMR (125 MHz, DMSO- $d_6$ )  $\delta$  172.7, 167.7, 163.6, 151.6, 136.7, 132.8, 131.2 (x2), 120.7 (x2), 100.5, 98.6 (1 missing quaternary carbon); HRMS calc. for C<sub>13</sub>H<sub>11</sub>NO<sub>4</sub> [M+H]<sup>+</sup> 246.0688; found 246.1010

**6-chloro-N,9-bis(tetrahydro-2H-pyran-2-yl)-9H-purin-2-amine (118)**



To a solution of 6-chloro-2-aminopurine (1g, 2.95 mmol) in anhydrous DMF (50 mL) was added 3,4-dihydropyran (2.7 mL, 29.5 mmol) followed by the addition of 5.7 M aqueous hydrochloric acid (50  $\mu$ L). The mixture was stirred at 60°C for 2 hours and then was evaporated *in vacuo*. Chromatography (SP4, silica, EtOAc/Petrol (4:6)) to give a yellow oil that crystallized upon freezing (1.55 g, 78%).  $R_f$  0.44 (100% EtOAc); m.p. 171-174 °C; IR  $\nu_{\text{max}}/\text{cm}^{-1}$  3269, 2940, 2855, 1716, 1614;  $^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ )  $\delta$  8.42 (s, 1H), 8.14 (br s, 1H), 5.53 (d,  $J$  = 10.2 Hz, 1H), 5.20-5.12 (m, 1H), 4.02 (d,  $J$  = 11.1, 1H), 3.84 (d,  $J$  = 11.1 Hz, 1H), 3.46 (td,  $J$  = 2.8, 11.1 Hz, 1 H), 2.21-2.35 (m, 1H), 2.03-1.34 (m, 11H);  $^{13}\text{C}$  NMR (125 MHz,  $(\text{CD}_3)_2\text{SO}$ )  $\delta$  158.1, 153.7, 150.0, 142.3, 124.7, 81.6, 80.4, 68.2, 66.0, 30.7, 30.1, 30.0, 25.4, 24.9, 22.9; HRMS calc. for  $\text{C}_{15}\text{H}_{21}\text{ClN}_4\text{O}_2$   $[\text{M} + \text{H}]^+$  338.1378; found 338.1378

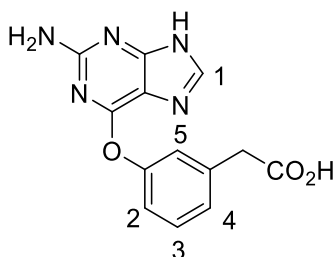
**Methyl 2-(3-((9-(tetrahydro-2H-pyran-2-yl)-2-((tetrahydro-2H-pyran-2-yl)amino)-9H-purin-6-yl)oxy)phenyl)acetate (119)**



DABCO (88 mg, 0.59 mmol) was added to a mixture of 6-chloro-*N*,9-*bis*(tetrahydro-2H-pyran-2-yl)-9H-purin-2-amine (100 mg, 0.30 mmol) and anhydrous DMSO (3 mL). The mixture was stirred at room temperature for 3 h. The mixture was added to a stirred suspension of DBU (49  $\mu$ L, 0.33 mmol) and methyl 2-(3-hydroxyphenyl)acetate (98 mg, 0.40 mmol) in anhydrous DMF (2 mL). The mixture was stirred at 65 °C for 72 h. The mixture was treated with aqueous saturated sodium bicarbonate (8 mL) and was extracted three times

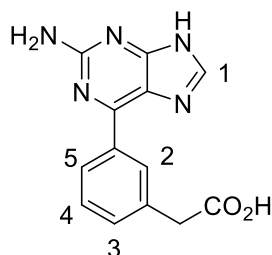
with EtOAc and the combined extracts washed three times with brine, dried over sodium sulfate, filtered, and concentrated *in vacuo* to give the crude product (145 mg) which was used as such in the next step without further purification.

### 2-(3-((2-amino-9H-purin-6-yl)oxy)phenyl)acetic acid (120)



The crude methyl 2-(3-((9-(tetrahydro-2H-pyran-2-yl)-2-((tetrahydro-2H-pyran-2-yl)amino)-9H-purin-6-yl)oxy)phenyl)acetate (142 mg, 0.3 mmol) was dissolved in THF (5 mL) and 2M aqueous HCl (2.5 mL) and stirred at room temperature for 4 h. The mixture was evaporated *in vacuo* purified by chromatography (SP4, C<sub>18</sub> silica, H<sub>2</sub>O:CH<sub>3</sub>CN:NH<sub>4</sub>OH 70:30:0.1) gave a white powder (8 mg, 10%). *R<sub>f</sub>* 0.18 (DCM/MeOH (9:1)); m.p. 284-285 °C; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 1233, 1265, 1376, 1568 cm<sup>-1</sup>; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  7.97 (1H, s, H-1), 7.38 (1H, t, *J* = 8.1 Hz, H-3), 7.27-7.21 (2H, m, H-2, H-5), 7.11 (1H, dd, *J* = 8.1, 1.1 Hz, H-4), 3.60 (2H, s, -NH<sub>2</sub>); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  173.2, 160.22, 160.18, 153.0, 138.2, 129.6, 126.5, 122.8, 120.2, 41.9; HRMS calc. for C<sub>15</sub>H<sub>14</sub>N<sub>3</sub>O<sub>3</sub> [M + H]<sup>+</sup> 286.0935; found 286.0931

### 2-(3-(2-amino-9H-purin-6-yl)phenyl)acetic acid (123)

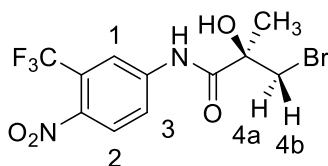


A mixture of 6-chloro-*N*,9-*bis*(tetrahydro-2H-pyran-2-yl)-9H-purin-2-amine (100 mg, 0.3 mmol), potassium carbonate (102 mg, 0.74 mmol), ethyl 2-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)acetate (86 mg, 0.3 mmol), tetrakis(triphenylphosphine)palladium(0) (17 mg, 0.02 mmol), anhydrous DMF (1.5 mL) and deionized water (5 mL) was stirred at 140 °C for 30 minutes under MW irradiation under N<sub>2</sub>. The mixture was treated with deionized water (5 mL) and extracted EtOAc (x3) and the

combined extracts washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered, and evaporated *in vacuo*. The crude residue was dissolved in THF (5 mL) and 2 M aqueous HCl (2.5 mL) and stirred at room temperature for 4 h. The mixture was neutralized with 2 M aqueous NaOH, extracted three times with EtOAc, and the combined extracts washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered, and concentrated *in vacuo* to give a yellow oil which was triturated with anhydrous methanol (5 mL) to afford a white powder (5 mg, 6%). *R*<sub>f</sub> 0.2 (DCM/MeOH (9:1)); m.p. 329–330 °C; IR  $\nu_{\text{max}}/\text{cm}^{-1}$  3380, 3163, 3041, 2113, 1911, 638; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  8.58 (d, *J* = 7.4 Hz, H-5), 8.50 (br s, H-1), 8.10 (1H, s, H-2), 7.48 (1H, t, *J* = 7.4 Hz, H-4), 7.40 (1H, d, *J* = 7.4 Hz, H-3), 3.69 (2H, s, -CH<sub>2</sub>-); <sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  174.8, 156.8, 148.6, 137.8, 135.5, 131.7, 130.6, 129.5, 41.5; HRMS calc. for C<sub>13</sub>H<sub>12</sub>N<sub>5</sub>O<sub>2</sub> [M + NH<sub>4</sub>]<sup>+</sup> 287.3025; found 287.1240

## 25.7 Synthesis of flutamide precursor<sup>297</sup>

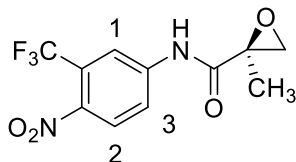
### (*R*)-3-Bromo-2-hydroxy-2-methyl-*N*-(4-nitro-3-(trifluoromethyl)phenyl)propanamide<sup>298</sup> (125)



To a solution of (2*R*)-3-bromo-2-hydroxy-2-methylpropionic acid (0.1 g, 0.49 mmol) in anhydrous THF (2 mL), SOCl<sub>2</sub> (0.04 mL, 0.49 mmol) was added at -5 °C under N<sub>2</sub> atmosphere. The mixture was stirred at 0 °C for 2 hours and then, 4-nitro-3-(trifluoromethyl)aniline in THF (2 mL) was added. The mixture was allowed to warm to room temperature slowly and stirred for 24 hours and then was evaporated *in vacuo*. The residue was taken up with DCM (10 mL), treated with saturated aqueous solution of NaHCO<sub>3</sub>, extracted with petrol (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, silica, EtOAc in Petrol, eluting 5–25%) gave a pale yellow oil (60 mg, 33%). *R*<sub>f</sub> 0.45 (EtOAc/Petrol (1:2)), mp 105–106 °C; <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  10.59 (1H, s, -NH-C(=O)), 8.56 (1H, d, *J* = 2.1 Hz, H-1), 8.36 (1H, dd, *J* = 9.0, 2.2 Hz, H-3), 8.21 (1H, d, *J* = 9.0 Hz, H-2), 6.41 (1H, s, -OH), 3.84 (1H, d, *J* = 10.4 Hz, H-4a), 3.60 (1H, d, *J* = 10.4 Hz, H-4b), 1.50 (s, 3H); (lit., *R*<sub>f</sub> 0.63 (EtOAc/hexanes (1:1)); m.p. 101–102 °C. <sup>1</sup>H NMR (DMSO-

$d_6$ ) 10.58 (s, N-H), 8.55 (1H, d,  $J = 2.0$  Hz), 8.35 (1H, dd,  $J = 2.0, 9.0$  Hz), 8.19 (1H, d,  $J = 9.0$  Hz), 6.41 (1H, s), 3.82 (1H, d,  $J = 10.5$  Hz), 3.58 (1H, d,  $J = 10.5$  Hz), 1.49 (3H, s, CH<sub>3</sub>))

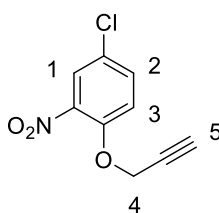
**(S)-2-Methyl-N-(4-nitro-3-(trifluoromethyl)phenyl)oxirane-2-carboxamide<sup>299</sup>(126)**



To a solution of (R)-3-bromo-2-hydroxy-2-methyl-N-(4-nitro-3-(trifluoromethyl)phenyl)propanamide (0.23 g, 0.6 mmol) in anhydrous acetone (5 mL) K<sub>2</sub>CO<sub>3</sub> (171 mg, 1.24 mmol) was added at room temperature. The mixture was stirred at 60°C for 6 hours and then was treated with water, extracted with EtOAc (x3), the combined organic extracts were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, C<sub>18</sub>, CH<sub>3</sub>CN [0.1% NH<sub>3</sub>] in water [0.1% NH<sub>3</sub>], eluting 10-80%) gave a pale yellow oil (320 mg, 68%). mp 88-89 °C; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 7.95 (1H, d,  $J = 9.0$  Hz, H-1), 7.03 (1H, d,  $J = 2.5$  Hz, H-2), 6.80 (1H, dd,  $J = 9.0, 2.5$  Hz, H-3), 2.65 (2H, s, -CH<sub>2</sub>-), 2.20 (3H, s, -CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>) δ 154.0, 135.3, 128.8, 123.6, 121.5, 114.2, 69.2, 54.6, 28.2 (2 missing carbon signals); (lit., mp 83-85 °C)

**25.8 Synthesis of (132)**

**4-Chloro-2-nitro-1-(prop-2-yn-1-yloxy)benzene<sup>300</sup>(129)**

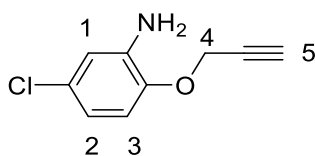


To a solution of 4-chloro-2-nitrophenol (0.5 g, 2.9 mmol) and K<sub>2</sub>CO<sub>3</sub> (0.6 g, 4.3 mmol) in anhydrous DMF (10 mL) (80%) propargyl bromide (0.42 mL, 3.8 mmol) was added dropwise at 0°C under N<sub>2</sub> atmosphere. The mixture was allowed to warm up to room temperature slowly and stirred for 16 hours, then was evaporated *in vacuo*. The residue was dissolved in water (10 mL), extracted with DCM (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a brown oil (592 mg, 97%) which was used in the next step without further purification. mp 81-82 °C;



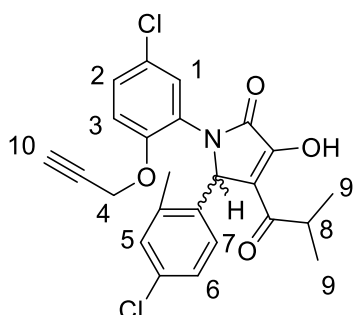
LCMS (ESI)<sup>+</sup>  $m/z$  = 210.6 [M + H]<sup>+</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm: 219.4; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3296, 2136, 2088, 1910, 1773, 1608, 1571, 1511; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  7.90 (1H, d,  $J$  = 2.6 Hz, H-1), 7.64 (1H, dd,  $J$  = 9.0, 2.6 Hz, H-2), 7.42 (1H, d,  $J$  = 9.0 Hz, H-3), 4.95 (2H, d,  $J$  = 2.4 Hz, -CH<sub>2</sub>-4), 3.12 (1H, t,  $J$  = 2.4 Hz, H-5); <sup>13</sup>C NMR (126 MHz, Chloroform-*d*)  $\delta$  150.5, 134.5, 126.6, 126.0, 118.4, 78.53, 78.52, 58.4 (1 missing carbon signal); not found by HRMS; (lit., mp 75-77 °C; <sup>1</sup>H NMR (300 MHz, Chloroform-*d*)  $\delta$  7.84 (1H, d,  $J$  = 2.6 Hz), 7.52 (1H, dd,  $J$  = 9.0, 2.6 Hz), 7.23 (1H, d,  $J$  = 9.0 Hz), 4.84 (2H, d,  $J$  = 2.4 Hz), 2.61 (1H, t,  $J$  = 2.4 Hz); <sup>13</sup>C NMR (75 MHz, Chloroform-*d*)  $\delta$  149.5, 140.5, 133.8, 126.6, 125.6, 117.1, 77.54, 76.8, 57.6)

### 5-Chloro-2-(prop-2-yn-1-yloxy)aniline (130)



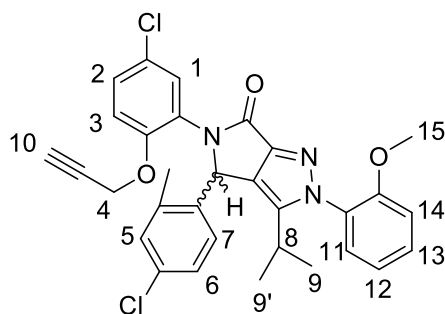
To a solution of 4-chloro-2-nitro-1-(prop-2-yn-1-yloxy)benzene (6 g, 28.4 mmol) in a mixture of dry ethanol (25 mL) and conc. HCl (2 mL) SnCl<sub>2</sub>·H<sub>2</sub>O (16 g, 70.9 mmol) was added portionwise at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at 80°C for 16 hours and then allowed to cool down to room temperature. 2M aqueous NaOH solution was added and the pH adjusted to 12. The mixture was extracted with EtOAc (x3), the combined organic extracts were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil which was filtered through an ionic exchange cartridge (SCX-2), washed with MeOH and subsequently with MeOH [0.1% NH<sub>3</sub>] affording a white solid (1.2 g, 23%) which was used in the next step without further purification. LCMS (ESI)<sup>+</sup>  $m/z$  = 182.3 [M + H]<sup>+</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm: 243.6, 295.4; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3427, 3390, 3274, 2114, 1610, 1492, 1445; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  6.89 (1H, d,  $J$  = 8.5 Hz, H-1), 6.75 (1H, d,  $J$  = 2.5 Hz, H-3), 6.60 (1H, dd,  $J$  = 8.5, 2.5 Hz, H-2), 4.75 (2H, d,  $J$  = 2.4 Hz, H-4), 2.98 (1H, t,  $J$  = 2.4 Hz, H-5); <sup>13</sup>C NMR (126 MHz, Chloroform-*d*)  $\delta$  145.4, 140.3, 128.0, 117.8, 115.8, 114.8, 79.8, 76.9, 57.4; HRMS calc. for C<sub>9</sub>H<sub>8</sub>ClNO [M+H]<sup>+</sup> 182.0367, found 182.0371 (lit.,<sup>301</sup> IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3472, 3380, 3291, 2122, 1614, 1504; <sup>1</sup>H NMR (300 MHz, Chloroform-*d*)  $\delta$  6.78 (1H, d,  $J$  = 8.5 Hz), 6.68- 6.61 (2H, m), 4.65 (2H, d,  $J$  = 2.4 Hz), 3.85 (2H, br s), 2.52 (1H, t,  $J$  = 2.4 Hz); <sup>13</sup>C NMR (75 MHz, Chloroform-*d*)  $\delta$  143.7, 138.0, 127.1, 117.5, 115.0, 113.5, 78.4, 75.8, 56.6)

**1-(5-Chloro-2-(prop-2-yn-1-yloxy)phenyl)-5-(4-chloro-2-methylphenyl)-3-hydroxy-4-isobutyryl-1,5-dihydro-2-*H*-pyrrol-2-one (131)**



A solution of 5-chloro-2-(prop-2-yn-1-yloxy)aniline (0.51 g, 2.8 mmol), 4-chloro-2-methyl benzaldehyde (0.43 g, 2.8 mmol) and ethyl isobutyrylacetate (0.52 g, 2.8 mmol) in acetic acid (5 mL) was stirred at 110°C for 24 hours. The solvent was evaporated *in vacuo* and the residue was co-evaporated with toluene affording a pale yellow oil. Chromatography (SP4, EtOAc in Petrol, eluting 35-90%) gave a brown oil (0.4 g, 31%).  $R_f$  0.44 (silica, EtOAc in Petrol (6:4)); LCMS (ESI)<sup>+</sup>  $m/z$  = 457.1 [M + H]<sup>+</sup>; UV  $\lambda_{max}$  (EtOH)/nm: 253.0; IR  $\nu_{max}/cm^{-1}$  2972, 2933, 1699, 1650, 1493, 1416; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  7.20 (1H, dd,  $J$  = 8.9, 2.6 Hz, H-2), 7.06 (1H, d,  $J$  = 8.9 Hz, H-3), 7.02 (1H, dd,  $J$  = 8.4, 2.1 Hz, H-6), 6.96 (3H, m, H-1, H-5, H-7), 6.03 (1H, s, -H (chiral centre)), 4.71 (2H, t,  $J$  = 2.3 Hz, -CH<sub>2</sub>-4), 3.55 (1H, p,  $J$  = 6.8 Hz, H-8), 2.99 (1H, t,  $J$  = 2.3 Hz, H-10), 2.32 (3H, s, -CH<sub>3</sub>), 1.01 (3H, d,  $J$  = 6.8 Hz, -CH<sub>3</sub>-9), 0.95 (3H, d,  $J$  = 6.8 Hz, -CH<sub>3</sub>-9); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  201.7, 168.6, 158.1, 153.6, 141.2, 135.6, 134.0, 130.8, 130.6, 130.1, 128.2, 127.3, 127.1, 126.8, 120.6, 115.9, 78.8, 77.9, 58.9, 57.6, 39.0, 19.4, 19.0, 18.1; HRMS calc. for C<sub>24</sub>H<sub>21</sub>Cl<sub>2</sub>NO<sub>4</sub> [M+H]<sup>+</sup> 458.0920, found 458.1820

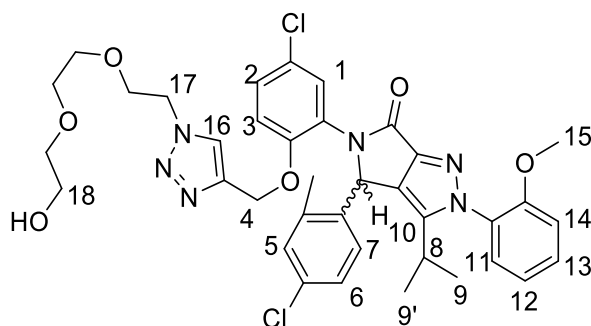
**5-(5-Chloro-2-(prop-2-yn-1-yloxy)phenyl)-4-(4-chloro-2-methylphenyl)-3-isopropyl-2-(2-methoxyphenyl)-4,5-dihydropyrrolo[3,4-*c*]pyrazol-6(2-*H*)-one (132)**



2-methoxy phenyl hydrazine hydrochloride (1.14 g, 6.6 mmol) was added portionwise to a stirred solution of 1-(5-chloro-2-(prop-2-yn-1-yloxy)phenyl)-5-(4-chloro-2-methylphenyl)-3-hydroxy-4-isobutyryl-1,5-dihydro-2H-pyrrol-2-one (0.5 g, 1.1 mmol) in dry EtOH (10 mL) and then a acetic acid (catalytic) was added to the mixture at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at 130°C for 16 hours and then was evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, EtOAc in Petrol, eluting 1-20%) gave a pale yellow oil (480 mg, 79%). *R<sub>f</sub>* 0.51 (silica, EtOAc in Petrol (6:4)); LCMS (ESI)<sup>-</sup> *m/z* = 559.1 [M - H]<sup>-</sup>; UV λ<sub>max</sub> (EtOH)/nm: 287.6, 264.6; IR ν<sub>max</sub>/cm<sup>-1</sup> 2963, 2926, 2558, 1705, 1485; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 7.54 (1H, dd, *J* = 7.8, 1.6 Hz, H-2), 7.52 – 7.47 (1H, m, H-11), 7.38 – 7.07 (7H, m, H-13, H-5, H-1, H-6, H-7, H-12, H-3), 7.05 (1H, d, *J* = 8.4 Hz, H-14), 6.46 (1H, s, -H (chiral centre)), 6.27 (0.5H, s, -H (chiral centre)), 4.79 (2H, t, *J* = 12.0 Hz, -CH<sub>2</sub>-4), 3.90 (3H, s, -CH<sub>3</sub>-15 (due to presence of conformers)), 3.87 (1.4H, s, -CH<sub>3</sub>-15 (due to presence of conformers)), 2.92 (1H, hept, *J* = 7.1 Hz, H-8), 2.35 (3H, s, -CH<sub>3</sub> (due to presence of conformers)), 2.05 (1.5H, s, (due to presence of conformers)), 1.14 (3H, t, *J* = 7.0 Hz, -CH<sub>3</sub>-9), 0.95 (3H, dd, *J* = 7.0 Hz, -CH<sub>3</sub>-9); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>) δ 158.1, 157.9, 154.0, 153.9, 152.7, 152.6, 152.2, 141.0, 140.7, 140.2, 138.8, 134.6, 133.9, 133.5, 132.8, 131.9, 131.7, 131.3, 130.3, 130.2, 130.1, 129.93, 129.91, 129.1, 129.0, 128.6, 128.5, 127.3, 126.8, 126.7, 126.6, 126.5, 126.4, 125.7, 125.50, 125.47, 120.27, 120.25, 114.6, 114.5, 112.2, 112.1, 77.60, 77.55, 76.59, 76.56, 62.7, 56.5, 56.2, 56.1, 55.2, 55.1, 27.53, 27.47, 21.2, 21.0, 20.6, 20.6, 17.9, 17.5

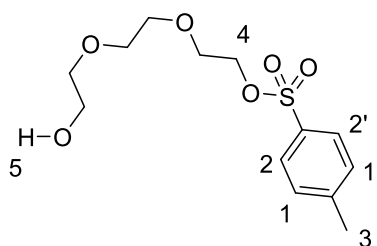
## 25.9 Synthesis of MDM2-based PROTACs for ARs

**5-(5-Chloro-2-((1-(2-(2-(2-hydroxyethoxy)ethoxy)ethyl)-1-*H*-1,2,3-triazol-4-yl)methoxy)phenyl)-4-(4-chloro-2-methylphenyl)-3-isopropyl-2-(2-methoxyphenyl)-4,5-dihydropyrrolo[3,4-*c*]pyrazol-6(2-*H*)-one (135)**



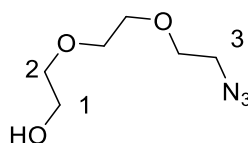
To a stirred solution of 5-(5-chloro-2-(prop-2-yn-1-yloxy)phenyl)-4-(4-chloro-2-methylphenyl)-3-isopropyl-2-(2-methoxyphenyl)-4,5-dihydropyrrolo[3,4-c]pyrazol-6(2H)-one (10 mg, 0.02 mmol) and (+)-L-sodium ascorbate (1 mg, catalytic) and copper (I) iodide (catalytic) in anhydrous THF (1 mL) 2-(2-(2-azidoethoxy)ethoxy)ethan-1-ol (3 mg, 0.01 mmol) in anhydrous THF (0.5 mL) was added at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at 110°C for 1 hour under MW irradiation and then was evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, MeOH in DCM, eluting 1-7%) gave a yellow oil (10 mg, 77%). R<sub>f</sub> 0.42 (silica, MeOH in DCM (1:9)); LCMS (ESI)<sup>-</sup> *m/z* = 734.6 [M - H]<sup>-</sup>; UV λ<sub>max</sub> (EtOH)/nm: 287.2, 265.4; IR ν<sub>max</sub>/cm<sup>-1</sup> 3414, 2924, 2869, 2101, 1908, 1704, 1596, 1559, 1493; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 8.07 (0.4 H, s, H-16 (due to presence of conformers)), 8.03 (0.6 H, s, H-16 (due to presence of conformers)), 7.55 – 7.42 (2H, m, H-1, H-11), 7.42 – 6.94 (7H, m, H-13, H-5, H-2, H-6, H-7, H-12, H-3), 6.97 (1H, d, *J* = 8.4 Hz, H-14), 6.43 (0.6 H, s, -H-10 (chiral centre)), 6.29 (0.6 H, s, -H-10 (chiral centre)), 5.36 – 5.03 (2H, m, -CH<sub>2</sub>-17), 4.61 – 4.45 (2H, m, -CH<sub>2</sub>-4), 3.88 – 3.80 (2H, m, -CH<sub>2</sub>-18), 3.79 (3H, d, *J* = 11.8 Hz, -CH<sub>3</sub>-15), 3.65 – 3.58 (2H, m, -CH<sub>2</sub>-(linker)), 3.56 – 3.51 (2H, m, -CH<sub>2</sub>-(linker)), 3.51 – 3.48 (2H, m, -CH<sub>2</sub>-(linker)), 3.47 – 3.43 (2H, m, -CH<sub>2</sub>-(linker)), 2.88 (1H, heptd, *J* = 7.0 Hz, -H-8), 2.32 (1.5 H, s, -CH<sub>3</sub> (due to presence of conformers)), 1.95 (1.2 H, s, -CH<sub>3</sub> (due to presence of conformers)), 1.08 (3H, dd, *J* = 7.0 Hz, -CH<sub>3</sub>-9 (due to presence of conformers)), 0.89 (3H, dd, *J* = 7.0 Hz, -CH<sub>3</sub>-9 (due to presence of conformers)); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>) δ 159.4, 159.2, 155.3, 154.5, 153.9, 153.5, 143.9, 142.3, 142.0, 140.2, 138.9, 135.9, 135.2, 134.9, 134.3, 133.4, 133.0, 132.9, 131.6, 131.5, 131.3, 130.4, 129.92, 129.88, 129.2, 128.7, 128.2, 128.1, 128.0, 127.9, 127.9, 127.2, 126.6, 126.6, 126.4, 126.3, 121.6, 116.2, 116.1, 113.6, 113.5, 73.6, 71.33, 71.32, 70.2, 63.6, 63.3, 63.1, 62.2, 57.7, 56.6, 56.5, 51.4, 28.9, 28.8, 22.6, 22.4, 22.0, 21.9, 21.5, 19.2, 18.7

## 2-(2-(2-Hydroxyethoxy)ethoxy)ethyl-4-methylbenzenesulfonate<sup>302</sup> (133)



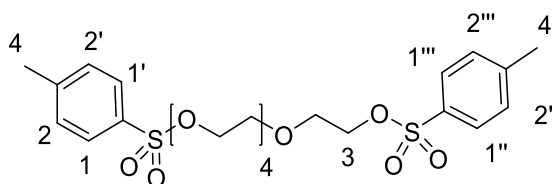
To a solution of triethyleneglycol (7.9 g, 53 mmoles) and triethylamine (1.1 mL, 7.9 mmoles) in anhydrous DCM (10 mL) was added *para*-toluenesulfonyl chloride (1 g, 5.3 mmoles) portionwise at 0°C under N<sub>2</sub> atmosphere. The mixture was allowed to warm to room temperature slowly and stirred for 16 hours and then was washed with saturated aqueous solution of NaHCO<sub>3</sub>, brine, the organic layer separated, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a pale yellow oil (1.1 g, 69%) which was used in the next step without further purification. <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 7.79 (2H, d, *J* = 8.2 Hz, H-2, H-2'), 7.49 (2H, d, *J* = 8.2 Hz, H-1, H-1'), 4.57 (1H, br s, -OH-5), 4.16 – 4.06 (2H, m, -CH<sub>2</sub>-4), 3.62 – 3.54 (2H, m, -CH<sub>2</sub>-(linker)), 3.50 – 3.33 (8H, m, -CH<sub>2</sub>-(linker)), 2.43 (3H, s, -CH<sub>3</sub>-3); <sup>13</sup>C NMR (126 MHz, DMSO-*d*<sub>6</sub>) δ 145.4, 132.9, 130.6 (x2), 128.1 (x2), 72.8, 70.5, 70.2, 70.1, 68.4, 60.7, 21.6; (lit., <sup>1</sup>H NMR (300 MHz, Chloroform-*d*) δ 7.78 (2H, d, *J* = 8.2 Hz), 7.33 (2H, d, *J* = 8.2 Hz), 4.14-4.16 (2H, m), 3.53-3.70 (10H, m), 2.51 (1H, s, OH), 2.43 (3H, s); <sup>13</sup>C NMR (75 MHz, Chloroform-*d*) δ 128.3, 72.9, 71.1, 70.6, 69.6, 69.0, 62.0, 22.0)

#### 2-(2-(2-Azidoethoxy)ethoxy)ethan-1-ol<sup>303</sup>(134)



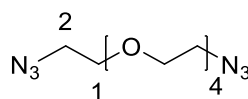
To a solution of 2-(2-(2-hydroxyethoxy)ethoxy)ethyl 4-methylbenzenesulfonate (1.1 g, 3.8 mmoles) in anhydrous DMF (20 mL) was added NaN<sub>3</sub> (0.31 g, 4.8 mmoles) at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at 100°C for 20 hours and then was treated with brine (20 mL), extracted with EtOAc (x3), the combined organic extracts were washed with brine (x2), dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a colourless oil. Chromatography (SP4, EtOAc in Petrol, eluting 40-90%) gave a colourless oil (0.4 g, 67%). <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 3.77-3.72 (2H, m, -CH<sub>2</sub>-1), 3.68 (6H, s, -CH<sub>2</sub>-(linker)), 3.64-3.60 (2H, m, -CH<sub>2</sub>-2), 3.43-3.37 (2H, m, -CH<sub>2</sub>-3); <sup>13</sup>C NMR (125 MHz, Chloroform-*d*) δ 72.5, 70.6, 70.4, 70.0, 61.7, 50.6; (lit., <sup>1</sup>H NMR (250 MHz, Chloroform-*d*) δ: 3.77-3.60 (10H, m), 3.41 (2H, t, *J* = 5.0 Hz); <sup>13</sup>C NMR (63 MHz, Chloroform-*d*) δ: 72.2, 70.2, 69.9, 69.6, 61.2, 50.2)

**3, 6, 9, 12-Tetraoxatetradecane-1,14-di-yl bis(4-methylbenzenesulfonate)<sup>304</sup> (138) (n=5)**



To a stirred solution of 3,6,9,12-tetraoxatetradecane-1,14-diol (4.11 g, 17.2 mmol) in DCM (50 ml) was added *para*-toluenesulfonyl chloride (6.58g, 34.5 mmol) at 0°C under N<sub>2</sub> atmosphere. KOH (7.69 g, 137 mmol) was ground up and added cautiously to the mixture and stirred at 0°C for 3 hours. The mixture was treated with water (50 ml), extracted with DCM (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil (6.4 g, 68%). LCMS (ESI)<sup>+</sup> *m/z* = 547.1 [M + H]<sup>+</sup>; UV λ<sub>max</sub> (EtOH)/nm: 224.6; IR ν<sub>max</sub>/cm<sup>-1</sup> 2869, 1597, 1451; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 7.81 (4H, d, *J* = 8.2 Hz, H-1, H-1', H-1'', H-1'''), 7.46 (4H, d, *J* = 8.2 Hz, H-2, H-2', H-2'', H-2'''), 4.19 – 4.13 (4H, m, -CH<sub>2</sub>-3 (x2)), 3.68 – 3.65 (4H, m, -CH<sub>2</sub>-(linker)), 3.59 (4H, s, -CH<sub>2</sub>-(linker)), 3.59–3.54 (8H, s, -CH<sub>2</sub>-(linker)), 2.47 (6H, s, -CH<sub>3</sub>-4); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>) δ 145.1 (x2), 133.1 (x2), 129.7 (x4), 127.7 (x4), 70.2 (x4), 70.1 (x2), 69.6 (x2), 68.3 (x2), 20.3 (x2); (lit., <sup>1</sup>H-NMR (400 MHz, Chloroform-*d*) δ 7.79 (4H, d, *J* = 8.4 Hz), 7.34 (4H, d, *J* = 8.5 Hz), 4.17 – 4.14 (4H, m), 3.69 – 3.67 (4H, m), 3.60 – 3.58 (12H, m), 2.44 (6H, s); <sup>13</sup>C-NMR (100 MHz, Chloroform-*d*) δ 145, 133, 130, 128, 70.9, 70.8, 70.7, 69.4, 68.9, 21.8)

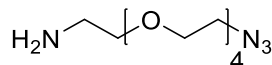
**1,14-Diazido-3,6,9,12-tetraoxatetradecane<sup>305</sup> (n=5) (139)**



To a stirred solution of 3,6,9,12-tetraoxatetradecane-1,14-diyl bis(4-methylbenzenesulfonate) (5.53 g, 10.1 mmol) and tetrabutylammonium iodide (0.187 g, 5 mol%) NaN<sub>3</sub> (2.63 g, 40.5 mmol) was added at room temperature under N<sub>2</sub> atmosphere. The mixture stirred at 80°C for 18 hours and then the precipitated formed was filtered through a sintered funnel and dried *in vacuo* affording a yellow oil (1.83 g, 63%). LCMS (ESI)<sup>+</sup> *m/z* = 289.2 [M + H]<sup>+</sup>; UV λ<sub>max</sub> (EtOH)/nm: 283.4; IR ν<sub>max</sub>/cm<sup>-1</sup> 2865, 2099, 1677; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 3.70 (4H, t, *J* = 4.8 Hz, -CH<sub>2</sub>-1), 3.68 (12H, s, -CH<sub>2</sub>-(linker)), 3.40 (4H, t, *J* = 4.8 Hz, CH<sub>2</sub>-2); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>) δ 70.3 (x4), 70.2 (x2), 69.8 (x2), 50.4 (x2); (lit., <sup>1</sup>H NMR

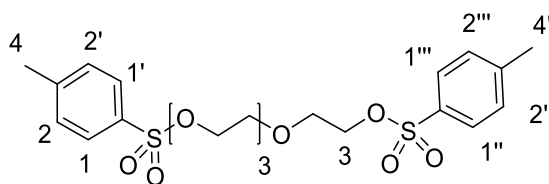
(300 MHz, Chloroform-*d*)  $\delta$  3.67 (12H, m), 3.38 (4H, t);  $^{13}\text{C}$  NMR (500 MHz, Chloroform-*d*)  $\delta$  70.97, 70.97, 70.29, 50.94)

**14-Azido-3,6,9,12-tetraoxatetradecan-1-amine (n=5) (140)**



To a solution of 1,14-diazido-3,6,9,12-tetraoxatetradecane (2.7 g, 9.4 mmol), Et<sub>2</sub>O (6 ml), EtOAc (6 ml) and 2M HCl (9.50 ml), Ph<sub>3</sub>P (2.45 g, 9.4 mmol) was added at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at room temperature for 24 hours and then was treated with water, washed with DCM (x3) and the aqueous layer separated, treated with 2M aqueous NaOH solution (10 ml) and the pH adjusted to 10. The aqueous layer was extracted with DCM (x3), the combined organic extracts were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil (1.2 g, 49%). LCMS (ESI)<sup>+</sup>  $m/z$  = 263.2 [M + H]<sup>+</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm: 272.4, 265.6; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 2863, 2097, 1594, 1452, 1347, 1285;  $^1\text{H}$  NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  3.71–3.65 (14H, m, -CH<sub>2</sub>-(linker)), 3.55 (2H, t,  $J$  = 5.3 Hz, -O-CH<sub>2</sub>-), 3.40 (2H, t,  $J$  = 5.3 Hz, NH<sub>2</sub>-CH<sub>2</sub>-), 2.81 (2H, t,  $J$  = 5.3 Hz, -CH<sub>2</sub>-N<sub>3</sub>);  $^{13}\text{C}$  NMR (126 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  73.8, 77.42, 72.37 (x4), 72.1, 71.9, 52.6, 42.8; (lit.,<sup>306</sup>  $^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$  3.66-3.61 (16 H, m), 3.48 (2 H, t,  $J$  = 5.2 Hz), 3.37 (2 H, t,  $J$  = 5.2 Hz), 2.84 (2 H, br);  $^{13}\text{C}$  NMR (100 MHz, Chloroform-*d*)  $\delta$  73.5, 70.7 (2X), 70.6 (3X), 70.3, 70.0, 50.8, 42.0)

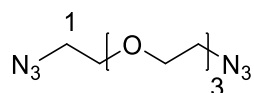
**((Oxybis(ethane-2,1-diyl))bis(oxy))bis(ethane-2,1-diyl) bis(4-methylbenzenesulfonate) (n=4) (138)**



To a stirred solution of 2,2'-((oxybis(ethane-2,1-diyl))bis(oxy))bis(ethan-1-ol) (10 g, 52 mmol) in anhydrous DCM (100 mL) *para*-toluenesulfonyl chloride (19.6 g, 103 mmol) was added at 0°C under N<sub>2</sub> atmosphere. KOH (23.1 g, 412 mmol) was ground up and added cautiously to the mixture and stirred at 0°C for 3 hours. The mixture was treated with water (50 mL), extracted with DCM (x3), the combined organic extracts were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a pale yellow oil (17.95 g, 69%). LCMS (ESI)<sup>+</sup>  $m/z$  = 503.1 [M + H]<sup>+</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm: 224.4; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 2871, 1597, 1451;  $^1\text{H}$  NMR

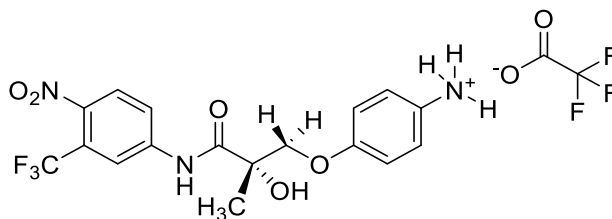
(500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  7.82 (4H, d,  $J$  = 8.3 Hz, H-1, H-1', H-1'', H-1'''), 7.46 (4H, d,  $J$  = 8.3 Hz, H-2, H-2', H-2'', H-2'''), 4.17 (4H, t,  $J$  = 4.6 Hz, -CH<sub>2</sub>-3), 3.67 (4H, t,  $J$  = 4.6 Hz, -O-CH<sub>2</sub>-(linker)), 3.55 (8H, s, -CH<sub>2</sub>-O-(linker)), 2.47 (6H, s, -CH<sub>3</sub>-4); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  147.3 (x2), 135.3 (x2), 131.9 (x4), 129.9 (x4), 72.4 (x2), 72.4 (x2), 71.8 (x2), 70.6 (x2), 22.4 (x2); (lit.,<sup>307</sup> <sup>1</sup>H NMR (400 MHz, Chloroform-*d*)  $\delta$  7.80 (4H, d,  $J$  = 8.2 Hz), 7.35 (4H, d,  $J$  = 8.2 Hz), 4.16 (4H, t,  $J$  = 4.7 Hz), 3.68 (4H, t,  $J$  = 4.7 Hz), 3.57 (8H, s), 2.45 (6H, s); <sup>13</sup>C NMR (400 MHz, Chloroform-*d*)  $\delta$  144.8, 133.0, 129.8, 128.0, 70.7, 70.5, 69.3, 68.7, 21.6)

### 1-Azido-2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethane (n=4) (139)



To a stirred solution of ((oxybis(ethane-2,1-diyl))bis(oxy))bis(ethane-2,1-diyl) bis(4-methylbenzenesulfonate) (16.33 g, 32.5 mmol) and tetrabutylammonium iodide (0.6 g, 5 mol%) in anhydrous DMF (60 mL) NaN<sub>3</sub> (8.45 g, 130 mmol) was added at room temperature under N<sub>2</sub> atmosphere. The mixture stirred at 80°C for 18 hours and then, the precipitate formed was filtered through a synered funnel and dried *in vacuo* affording a yellow oil (6.97 g, 88%). LCMS (ESI)<sup>+</sup>  $m/z$  = 245.1 [M + H]<sup>+</sup>; UV  $\lambda_{\max}$  (EtOH)/nm: 283.8; IR  $\nu_{\max}$ /cm<sup>-1</sup> 2866, 2091, 1677; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  3.71–3.67 (12H, m, -O-CH<sub>2</sub>-(linker)), 3.40 (4H, t,  $J$  = 4.90 Hz, -CH<sub>2</sub>-1); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  72.6 (x2), 72.4 (x2), 72.0 (x2), 52.6 (x2); (lit.,<sup>308</sup> <sup>1</sup>H NMR (400 MHz, Chloroform-*d*)  $\delta$  3.68 (12H, t,  $J$  = 6.7 Hz), 3.40 (4H, t,  $J$  = 6.7 Hz). <sup>13</sup>C NMR (100 MHz, Chloroform-*d*)  $\delta$  70.9 (x2), 70.2 (x4), 50.8 (x2))

### (S)-4-(2-Hydroxy-2-methyl-3-((4-nitro-3-(trifluoromethyl)phenyl)amino)-3-oxopropoxy)benzenaminium 2,2,2-trifluoroacetate (127)

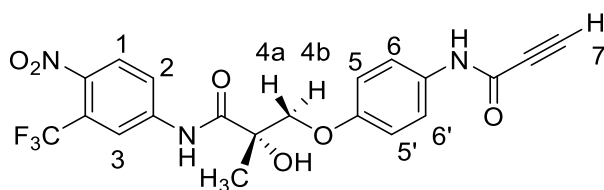


To a solution of (S)-2-methyl-N-(4-nitro-3-(trifluoromethyl)phenyl)oxirane-2-carboxamide (200 mg, 0.69 mmoles) in IPA (20 mL) and K<sub>2</sub>CO<sub>3</sub> (238 mg, 1.7 mmoles) N-Boc-4-



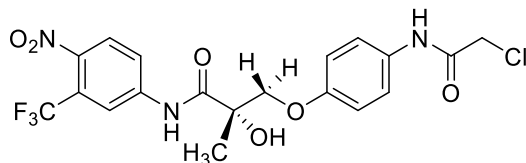
hydroxyaniline (144 mg, 0.69 mmoles) was added at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at 80°C for 4 hours and then treated with saturated aqueous solution of NaHCO<sub>3</sub>, extracted with DCM (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, MeOH in DCM, eluting 1-10%) gave a pale yellow oil (208 mg, 60%). LCMS (ESI)<sup>+</sup>  $m/z$  = 400.1 [M + H]<sup>+</sup> Then the oil was dissolved in TFA/DCM (1:1) (4 mL) and stirred at room temperature for 16 hours, then was evaporated *in vacuo* affording a white solid (90 mg, 54%) that was confirmed by LCMS and used in the next steps without further purification. LCMS (ESI)<sup>+</sup>  $m/z$  = 300.1 [M + H]<sup>+</sup>

**(S)-N-(4-(2-Hydroxy-2-methyl-3-((4-nitro-3-(trifluoromethyl)phenyl)amino)-3-oxopropoxy)phenyl)propiolamide (146)**



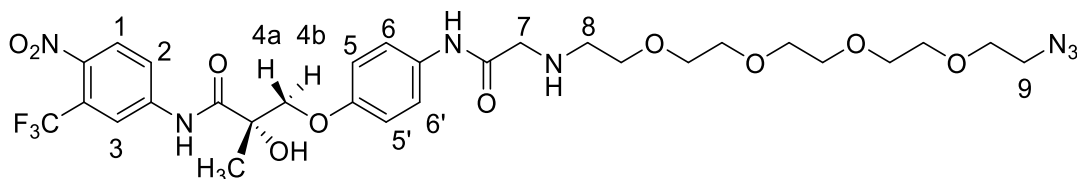
To a solution of (S)-3-(4-aminophenoxy)-2-hydroxy-2-methyl-N-(4-nitro-3-(trifluoromethyl)phenyl)propanamide (100 mg, 0.25 mmol) and propiolic acid (26 mg, 0.38 mmol) in anhydrous THF (2 ml) 4-dimethylaminopyridine (0.3 mg, 1 mol%) was added at room temperature. The mixture was stirred for 10 minutes at 0°C and then 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrogen chloride (73 mg, 0.38 mmol) was added. The mixture was stirred for 1 hour and then was diluted with DCM and washed with NH<sub>4</sub>Cl (x3). The organic layer was separated, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a brown oil (79 mg, 70%). LCMS (ESI)<sup>+</sup>  $m/z$  = 450.1 [M + H]<sup>+</sup>; UV  $\lambda_{\text{max}}$  (EtOH)/nm: 282.6; IR  $\nu_{\text{max}}$ /cm<sup>-1</sup> 3277, 2929, 2105, 1645, 1596, 1509, 1415; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  8.40 (1H, d,  $J$  = 2.3 Hz, H-3), 8.21 (1H, dd,  $J$  = 9.0, 2.3 Hz, H-2), 8.08 (1H, d,  $J$  = 9.0 Hz, H-1), 7.49 (2H, d,  $J$  = 9.1 Hz, H-6, H-6'), 6.95 (2H, d,  $J$  = 9.1 Hz, H-5, H-5'), 4.34 (1H, d,  $J$  = 9.5 Hz, H-4a), 4.05 (1H, d,  $J$  = 9.5 Hz, H-4b), 3.76 (1H, ddd,  $J$  = 6.6, 4.2, 2.5 Hz, -OH), 1.94 – 1.87 (1H, m, H-7), 1.56 (3H, s, -CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  210.9, 201.7, 177.0, 175.8, 128.8, 125.1, 123.7 (x2), 116.9 (x2), 77.4, 77.1, 76.0, 69.7, 31.5, 27.4, 24.0; HRMS calc. for C<sub>20</sub>H<sub>16</sub>F<sub>3</sub>N<sub>3</sub>O<sub>6</sub> [M+H]<sup>+</sup> 452.1064, found 452.1827

**(S)-3-(4-(2-Chloroacetamido)phenoxy)-2-hydroxy-2-methyl-N-(4-nitro-3-(trifluoromethyl)phenyl)propanamide (141)**



To a solution of (S)-3-(4-aminophenoxy)-2-hydroxy-2-methyl-N-(4-nitro-3-(trifluoromethyl)phenyl)propanamide (90 mg, 0.23 mmol) in THF (1.5 mL) and DIPEA (41  $\mu$ L, 0.23 mmol) chloroacetyl chloride (18  $\mu$ L, 0.23 mmol) was added at 0°C under N<sub>2</sub> atmosphere. The mixture was allowed to warm up to room temperature slowly and stirred under for 3 hours, then was treated with EtOAc and washed with 2M aqueous HCl solution and brine. The organic layer was separated, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil (107 mg, 99.99%) which was used without further purification. LCMS (ESI)<sup>+</sup>  $m/z$  = 475.1 [M + H]<sup>+</sup>

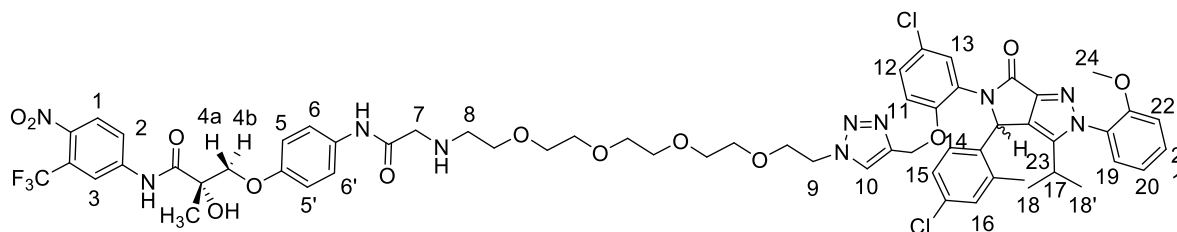
**(S)-17-Azido-N-(4-(2-hydroxy-2-methyl-3-((4-nitro-3-(trifluoromethyl)phenyl)amino)-3-oxopropoxy)phenyl)-6,9,12,15-tetraoxa-3azaheptadecanamide (142)**



To a solution of (S)-N-(4-(2-hydroxy-2-methyl-3-((4-nitro-3-(trifluoromethyl)phenyl)amino)-3-oxopropoxy)phenyl)propiolamide (80 mg, 0.17 mmol), 14-azido-3,6,9,12-tetraoxatetradecan-1-amine (53 mg, 0.2 mmol) in anhydrous DMF (2 mL), K<sub>2</sub>CO<sub>3</sub> (29.2 mg, 0.21 mmol) was added at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at 100°C for 3 hours and then was treated with saturated aqueous solution of NaHCO<sub>3</sub>, extracted with DCM (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, MeOH in DCM, eluting 1-10%) gave a yellow oil (41 mg, 35%). LCMS (ESI)<sup>-</sup>  $m/z$  = 700.3 [M - H]<sup>-</sup>; UV  $\lambda_{\max}$  (EtOH)/nm: 250.6; IR  $\nu_{\max}$ /cm<sup>-1</sup> 3296, 2868, 2100, 1672, 1595, 1509, 1414; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  8.40 (1H, d,  $J$  = 2.1 Hz, H-3), 8.20 (1H, dd,  $J$  = 8.9, 2.1 Hz, H-2), 8.07 (1H, d,  $J$  = 8.9 Hz, H-1), 7.50 (2H, d,  $J$  = 9.0 Hz, H-6), 6.94 (2H, d,  $J$  = 9.0 Hz, H-5), 4.34 (1H, d,  $J$  = 9.5 Hz, H-4a), 4.04 (1H, d,  $J$  = 9.5 Hz, H-4b), 3.76 (1H, t,  $J$  = 6.6 Hz, -OH), 3.69 – 3.58 (16H, m, -CH<sub>2</sub>-(linker)), 3.42 (2H, s, -CH<sub>2</sub>-7), 3.39 – 3.35 (2H,

m, -CH<sub>2</sub>-8), 2.89 – 2.80 (2H, m, -CH<sub>2</sub>-9), 1.56 (3H, s, -CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>) δ 210.9, 177.0, 172.9, 157.8, 144.8, 133.8, 128.8, 123.6 (x2), 120.5, 116.9 (x2), 77.4, 76.0, 72.5, 72.44, 72.39, 72.36, 72.14, 72.11, 71.9, 69.7, 54.1, 52.6, 50.7, 31.5, 24.0 (2 missing carbon signal); HRMS calc. for C<sub>29</sub>H<sub>38</sub>F<sub>3</sub>N<sub>7</sub>O<sub>10</sub> [M+H]<sup>+</sup> 702.2705, found 702.2452

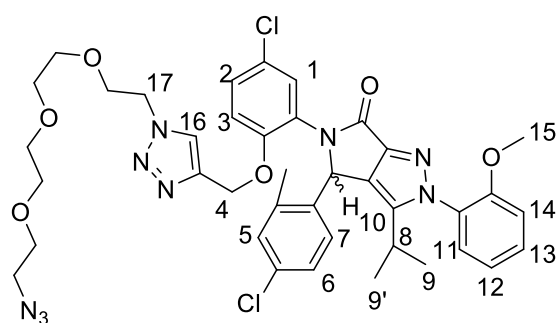
**(S)-17-(4-((4-Chloro-2-(4-(4-chloro-2-methylphenyl)-3-isopropyl-2-(2-methoxyphenyl)-6-oxo-2,6-dihydropyrrolo[3,4-c]pyrazol-5(4-*H*)-yl)phenoxy)methyl)-1-*H*-1,2,3-triazol-1-yl)-*N*-(4-(2-hydroxy-2-methyl-3-((4-nitro-3-(trifluoromethyl)phenyl)amino)-3-oxopropoxy)phenyl)-6,9,12,15-tetraoxa-3-azaheptadecanamide (143)**



To a stirred solution of 5-(5-chloro-2-(prop-2-yn-1-yloxy)phenyl)-4-(4-chloro-2-methylphenyl)-3-isopropyl-2-(2-methoxyphenyl)-4,5-dihydropyrrolo[3,4-c]pyrazol-6(2-*H*)-one (24 mg, 0.043 mmol), (+)-L-sodium ascorbate (catalytic) and 1mg of copper (II) sulfate pentahydrate (catalytic) in anhydrous THF (1 ml) tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (2 mg, 0.01 mmol) was added at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at 110°C for 1.5 hours under microwave irradiation. Then 1 mg of copper (II) sulfate pentahydrate was added and the reaction was re-subjected to the same conditions as described before, then the solvent was evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, MeOH in DCM, eluting 1-7%) gave a colourless oil (15 mg, 25%). *R*<sub>f</sub> 0.58 (silica, MeOH in DCM (1:99)); LCMS (ESI)<sup>−</sup> *m/z* = 1260.0 [M - H]<sup>−</sup>; UV λ<sub>max</sub> (EtOH)/nm: 285.5; IR ν<sub>max</sub>/cm<sup>−1</sup> 3308, 3128, 3072, 2920, 2869, 1697, 1595, 1509, 1457, 1413; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) δ 8.39 (1H, d, *J* = 2.1 Hz, H-3), 8.18 (1H, d, *J* = 8.9 Hz, H-2), 8.04 (2H, m, H-1, H-10), 7.55 – 7.48 (2H, m, H-13, H-19), 7.47 (2H, d, *J* = 8.9 Hz, H-6, H-6'), 7.42 – 6.96 (7H, m, H-21, H-15, H-12, H-15, H-14, 20, H-11), 6.91 (2H, d, *J* = 8.9 Hz, H-5, H-5'), 6.43 (0.6 H, d, *J* = 73.4 Hz, H-23, 6.29 (0.4 H, d, *J* = 73.4 Hz, H-23, 5.31 – 5.13 (2H, m, H-9), 4.45 – 4.58 (2H, m, -CH<sub>2</sub>-15), 4.33 (1H, d, *J* = 9.5 Hz, H-4a), 4.02 (1H, d, *J* = 9.5 Hz, H-4b), 3.83 (5H, m, -CH<sub>3</sub>-24, -CH<sub>2</sub>-7), 3.70 – 3.37 (16H, m, -CH<sub>2</sub>-(linker)), 2.90 (1H, pd, *J* = 7.0, 3.0 Hz, -H-17), 2.83 (2H, d, *J* = 5.0 Hz, H-8), 2.12 (d, *J* = 148.5 Hz, -CH<sub>3</sub>), 1.10 (3H, dd, *J* = 7.0, 1.7 Hz, H-18), 0.96 (2H, dd, *J* = 7.0 Hz, -CH<sub>3</sub>-18'), 0.88 (1H, dd,

$J = 7.0$  Hz,  $-\text{CH}_3\text{-18'}$ );  $^{13}\text{C}$  NMR (126 MHz, Methanol- $d_4$ )  $\delta$  177.0, 157.8, 156.2, 155.4, 154.8, 145.1, 141.1, 136.1, 135.8, 135.2, 134.3, 133.8, 132.4, 131.3, 130.8, 129.4, 128.8, 128.1, 127.5, 126.0, 125.1, 123.6 (x2), 122.5, 120.6, 117.0 (x2), 114.5, 77.4, 76.1, 71.1, 64.2, 58.6, 57.5, 53.8, 52.3, 29.7, 24.0, 23.3, 22.9, 20.0, 19.6; Analytical HPLC purity: 91.2% ( $\text{CH}_3\text{CN}$  [0.1% formic acid]); HRMS calc. for  $\text{C}_{60}\text{H}_{65}\text{Cl}_2\text{F}_3\text{N}_{10}\text{O}_{13}$ : 1261.4135, found 1261.4069

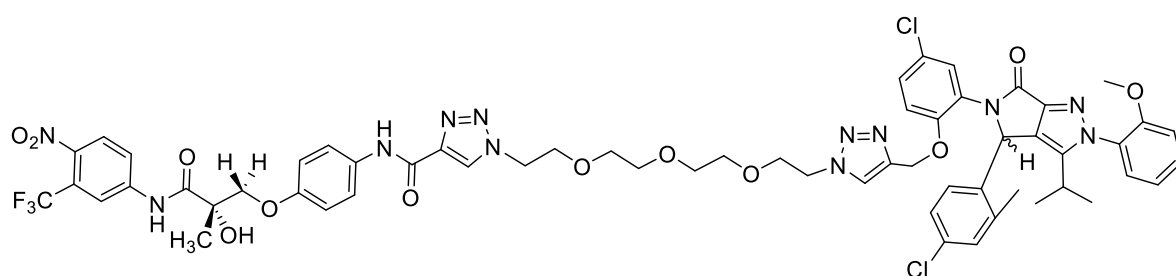
**5-(2-((1-(2-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)ethyl)-1-*H*-1,2,3-triazol-4-yl)methoxy)-5-chlorophenyl)-4-(4-chloro-2-methylphenyl)-3-isopropyl-2-(2-methoxyphenyl)-4,5-dihydropyrrolo[3,4-*c*]pyrazol-6(2-*H*)-one (147)**



To a stirred solution of 5-(5-chloro-2-(prop-2-yn-1-yloxy)phenyl)-4-(4-chloro-2-methylphenyl)-3-isopropyl-2-(2-methoxyphenyl)-4,5-dihydropyrrolo[3,4-*c*]pyrazol-6(2-*H*)-one (50 mg, 0.089 mmol) and 1-azido-2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethane (218 mg, 0.892 mmol), (+)-L-sodium ascorbate (catalytic) and copper (II) sulfate pentahydrate (catalytic) in anhydrous THF (1 ml) tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (5 mg, 0.01 mmol) was added at room temperature under  $\text{N}_2$  atmosphere. The mixture was stirred at  $110^\circ\text{C}$  for 1 hour under MW irradiation and then filtered through a sintered funnel and dried *in vacuo* affording a yellow oil (50 mg, 69%) which was used for the next step without further purification. LCMS (ESI) $^-$   $m/z = 803.3$  [ $\text{M} - \text{H}$ ] $^-$ ; UV  $\lambda_{\text{max}}$  (EtOH)/nm: 287.2, 265.8; IR  $\nu_{\text{max}}/\text{cm}^{-1}$  2867, 2097, 1710, 1595, 1559, 1493, 1460, 1413;  $^1\text{H}$  NMR (500 MHz, Methanol- $d_4$ )  $\delta$  8.09 (0.4H, s, H-16), 8.05 (0.6H, s, H-16), 7.57 – 7.46 (2H, m, H-1, H-11), 7.42 – 7.08 (7H, m, H-13, H-5, H-2, H-6, H-7, H-12, H-3), 7.00 (d,  $J = 8.3$  Hz, H-14), 6.44 (0.6H, s, H-10), 6.29 (0.4H, s, H-10) 5.34 – 5.17 (2H, m,  $-\text{CH}_2\text{-17}$ ), 4.64 – 4.51 (2H, m,  $-\text{CH}_2\text{-4}$ ), 3.93 – 3.80 (5H, m,  $-\text{CH}_2\text{-(linker)}$ ,  $-\text{CH}_3\text{-15}$ ), 3.64 – 3.47 (12H, m,  $-\text{CH}_2\text{-(linker)}$ ), 3.40 (2H, t,  $J = 4.9$  Hz,  $-\text{CH}_2\text{-17}$ ), 2.91 (1H, pd,  $J = 7.0$  Hz, H-8), 2.28-1.98 (3H, d,  $J = 149.91$  Hz,  $-\text{CH}_3$ ), 1.12 (3H, dd,  $J = 7.0$  Hz,  $-\text{CH}_3\text{-9}$ ), 0.93 (3H, dd,  $J = 7.0$  Hz,  $-\text{CH}_3\text{-9'}$ );  $^{13}\text{C}$  NMR (126 MHz, Methanol- $d_4$ )  $\delta$  158.0, 153.9, 153.2, 152.6, 152.1, 141.0, 140.6, 138.9,

134.5, 133.8, 133.6, 133.0, 132.0, 131.7, 131.5, 130.3, 130.1, 130.0, 129.0 (x2), 128.6, 128.5, 127.3, 126.9, 126.8, 126.7, 126.6, 126.5, 125.8, 125.3, 125.2, 125.2, 120.3 (x2), 114.8, 114.7, 112.2, 112.1, 70.3, 70.20, 70.17, 70.14, 70.09, 70.0, 69.74, 69.67, 68.9, 62.0, 61.8, 56.4, 55.24, 55.15, 50.39, 50.36, 50.1, 27.51, 27.45, 21.2, 21.0, 20.7, 20.6, 17.8, 17.4,

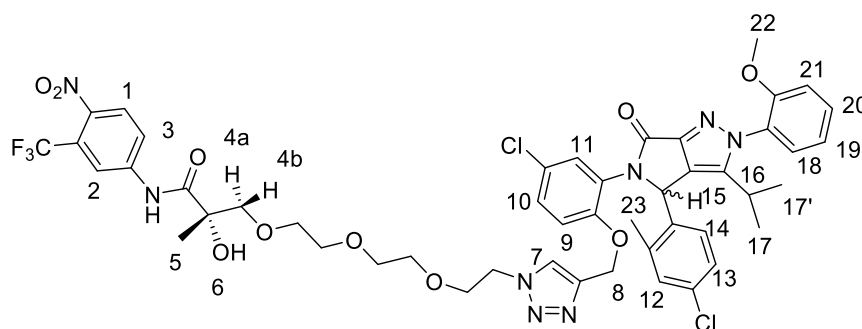
**1-(2-(2-(2-(2-(4-((4-Chloro-2-(4-(4-chloro-2-methylphenyl)-3-isopropyl-2-(2-methoxyphenyl)-6-oxo-2,6-dihydropyrrolo[3,4-c]pyrazol-5(4-*H*)-yl)phenoxy)methyl)-1-*H*-1,2,3-triazol-1-yl)ethoxy)ethoxy)ethoxy)ethyl)-*N*-(4-((*S*)-2-hydroxy-2-methyl-3-((4-nitro-3-(trifluoromethyl)phenyl)amino)-3-oxopropoxy)phenyl)-1-*H*-1,2,3-triazole-4-carboxamide (148)**



To a stirred solution of 5-(2-((1-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-1*H*-1,2,3-triazol-4-yl)methoxy)-5-chlorophenyl)-4-(4-chloro-2-methylphenyl)-3-isopropyl-2-(2-methoxyphenyl)-4,5-dihydropyrrolo[3,4-c]pyrazol-6(2*H*)-one (40 mg, 0.05 mmol), (*S*)-*N*-(4-(2-hydroxy-2-methyl-3-((4-nitro-3-(trifluoromethyl)phenyl)amino)-3-oxopropoxy)phenyl)propiolamide (22 mg, 0.05 mmol), (+)-*L*-sodium ascorbate (catalytic), copper (II) sulfate pentahydrate (catalytic) in THF (1 ml) tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (3 mg, 0.01 mmol) was added at room temperature under N<sub>2</sub> atmosphere. The mixture was stirred at 110°C for 1 hour under microwave irradiation and then was purified by Semi-Prep HPLC (SCX, Acetonitrile/Water +0.1% formic acid, 5/95 to 100/0) affording a dark brown oil (10 mg, 15%). LCMS (ESI)<sup>+</sup> *m/z* = 1255.4 [M + H]<sup>+</sup>; UV λ<sub>max</sub> (EtOH)/nm: 285.2; IR ν<sub>max</sub>/cm<sup>-1</sup> 3345, 3073, 2923, 2867, 2100, 1706, 1596, 1559, 1494, 1413; <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>) 8.47 (1H, d, *J* = 2.8 Hz, Ar-H), 8.14 (1H, d, *J* = 9.5 Hz, Ar-H), 8.02-7.97 (1H, m, Ar-H), 7.59 (1H, d, *J* = 8.9 Hz, Ar-H), 7.55-7.45 (2H, m, 2×Ar-H), 7.32-6.95 (10H, m, 10×Ar-H), 6.44-6.28 (1H, m, -C\*-H), 5.29-5.12 (2H, m, -O-CH<sub>2</sub>-CH<sub>2</sub>-N), 4.62-4.49 (2H, m, Ar-O-CH<sub>2</sub>-C-), 4.35 (1H, d, *J* = 9.5 Hz, -C(Me)(OH)-CH<sub>2</sub>-O-), 4.05 (1H, dd, *J* = 9.5, 1.1 Hz, -C(Me)(OH)-CH<sub>2</sub>-O-), 3.90-3.80 (5H, m, O-CH<sub>2</sub>-CH<sub>2</sub>-N, Ar-O-CH<sub>3</sub>), (3.63-3.44 (10H, m, 5×-CH<sub>2</sub>-), 3.38 (12H, s, 6×-CH<sub>2</sub>-), 2.91 (1H, h, *J* = 7.1 Hz, -CH(Me)<sub>2</sub>), 2.28-1.97 (3H, m, Ar-CH<sub>3</sub>), 1.56-1.32 (3H, m, -C\*-CH<sub>3</sub>), (1.13-1.08 (3H, m, -CH(CH<sub>3</sub>)<sub>2</sub>), 0.98-0.86 (3H, m, -CH-(CH<sub>3</sub>)<sub>2</sub>); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>) δ 177.0, 161.2,

158.1, 156.2, 155.4, 154.4, 145.1, 144.8, 142.9, 142.3, 141.1, 136.1, 135.8, 135.1, 134.3, 133.9, 133.5, 132.5, 132.4, 132.2, 131.3, 130.8, 129.5, 129.4, 129.1, 129.0, 128.82, 128.78, 128.0, 127.5, 127.3, 125.0, 124.4, 122.5, 117.1, 116.9, 114.5, 77.4, 76.0, 72.4, 72.34, 72.30, 72.26, 72.2, 71.9, 71.12, 71.06, 70.9, 64.2, 57.4, 57.3, 52.6, 52.4, 52.3, 52.3, 29.7, 24.0, 23.4, 23.2, 22.9, 22.8, 20.0, 19.6; Analytical HPLC purity: 100% (CH<sub>3</sub>CN [0.1% formic acid]), HRMS calc. for C<sub>59</sub>H<sub>59</sub>Cl<sub>2</sub>F<sub>3</sub>N<sub>12</sub>O<sub>12</sub>: 1254.3705, found 1254.3675

**(2S)-3-(2-(2-(2-(4-((4-chloro-2-(4-(4-chloro-2-methylphenyl)-3-isopropyl-2-(2-methoxyphenyl)-6-oxo-2,6-dihydropyrrolo[3,4-c]pyrazol-5(4H)-yl)phenoxy)methyl)-1H-1,2,3-triazol-1-yl)ethoxy)ethoxy)ethoxy)-2-hydroxy-2-methyl-N-(4-nitro-3-(trifluoromethyl)phenyl)propenamide (149)**



To a stirred solution of 5-(5-chloro-2-((1-(2-(2-(2-hydroxyethoxy)ethoxy)ethyl)-1H-1,2,3-triazol-4-yl)methoxy)phenyl)-4-(4-chloro-2-methylphenyl)-3-isopropyl-2-(2-methoxyphenyl)-4,5-dihydropyrrolo[3,4-c]pyrazol-6(2H)-one (0.1 g, 0.14 mmol) in anhydrous THF (5 mL) sodium hydride (60%) (3.6 mg, 0.15 mmol) was added at 0°C under N<sub>2</sub> atmosphere. The mixture was stirred at 0°C for 30 minutes and then (*S*)-2-Methyl-N-(4-nitro-3-(trifluoromethyl)phenyl)oxirane-2-carboxamide (42 mg, 0.14 mmol) at 0°C under N<sub>2</sub> atmosphere. Then the mixture was stirred at room temperature for 5 hours and then treated with saturated aqueous solution of NaHCO<sub>3</sub>, extracted with DCM (x3), the combined organic extracts were washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated *in vacuo* affording a yellow oil. Chromatography (SP4, MeOH in DCM, eluting 1-10%) gave a pale white solid (7 mg, 5%). LCMS (ESI)<sup>+</sup> *m/z* = 1025.3 [M + H]<sup>+</sup>; UV λ<sub>max</sub> (EtOH)/nm: 260.8; IR ν<sub>max</sub>/cm<sup>-1</sup> 728, 800, 834, 919, 1043, 1139, 1249, 1321, 1418, 1476, 1510, 1594, 1672, 2866, 2923, 3180, 3242, 3317; <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 9.74 (1H, br s, -NH(C=O)), 8.20 (1H, d, *J* = 2.0 Hz, H-2), 8.08 (1H, dd, *J* = 8.9, 2.2 Hz, H-3), 7.96 (1H, d, *J* = 8.9 Hz, H-1), 7.90 (1H, s, H-11), 7.59 (1H, dd, *J* = 7.8, 1.7 Hz, H-18), 7.44 (1H, d, *J* = 8.2 Hz, H-14),

7.38 (1H, td,  $J = 7.9, 1.7$  Hz, H-20), 7.24 (1H, d,  $J = 2.0$  Hz, H-12), 7.20 (1H, dd,  $J = 8.3, 2.1$  Hz, H-13), 7.09 (1H, td,  $J = 7.7, 1.2$  Hz, H-19), 7.01 – 6.93 (3H, m, H-7, H-21, H-9, H-10), 5.35 (2H, s, -CH<sub>2</sub>-8), 4.48 (2H, qdd,  $J = 14.4, 6.1, 4.0$  Hz, -CH<sub>2</sub>-), 3.96 (1H, d,  $J = 9.8$  Hz, H-4a), 3.85 (2H, dq,  $J = 10.5, 6.5$  Hz, -CH<sub>2</sub>-), 3.66 (s, 3H, s, -O-CH<sub>3</sub>-22), 3.64 – 3.53 (2H, m, -CH<sub>2</sub>-), 3.50 – 3.44 (4H, m, 2x-CH<sub>2</sub>-), 3.43 (d,  $J = 9.7$  Hz, H-4b), 2.67 (1H, hept,  $J = 6.9$  Hz, H-16), 2.46 (3H, s, H-23), 1.40 (3H, s, -CH<sub>3</sub>-5), 1.19 (6H, d,  $J = 6.8$  Hz, 2x (-CH<sub>3</sub>), H-17, H-17'); <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  194.0, 174.2, 158.3, 157.4, 152.6, 145.5, 143.3, 142.1, 140.0, 139.9, 137.9, 131.7, 131.6, 130.1, 129.1, 128.6, 127.1, 126.7, 125.9, 124.3, 124.0, 122.1, 121.2, 120.4, 119.6, 118.4, 113.1, 111.9, 76.0, 75.9, 71.0, 70.4, 70.3, 70.2, 69.2, 63.0, 55.6, 50.4, 27.0, 22.9, 22.4, 22.3, 20.3; Analytical HPLC purity: 100% (CH<sub>3</sub>CN [0.1% formic acid]), HRMS calc. for C<sub>59</sub>H<sub>59</sub>Cl<sub>2</sub>F<sub>3</sub>N<sub>12</sub>O<sub>12</sub>: 1024.2901, found 1024.3075

- 
- <sup>1</sup> <https://www.who.int/en/news-room/fact-sheets/detail/cancer>
- <sup>2</sup> <https://www.cancer.gov/about-cancer/understanding/what-is-cancer>
- <sup>3</sup> Knudson, A.G., Two genetic hits (more or less) to cancer, *Nature*, **2001**, 1, 157-162
- <sup>4</sup> <http://www.searo.who.int/publications/bookstore/documents/9283204298/en/>
- <sup>5</sup> Doll, R., Peto, R., The Causes of Cancer: quantitative estimates of avoidable risks of cancer in United States today, *J Natl Cancer Inst.*, **1981**, 66 (6), 1191-308
- <sup>6</sup> <https://www.iarc.fr/> (International Agency for Cancer Research – WHO)
- <sup>7</sup> <https://www.cancerresearchuk.org/health-professional/cancer-statistics/worldwide-cancer/mortality#heading-Zero>
- <sup>8</sup> <https://ourworldindata.org/cancer>
- <sup>9</sup> Hanahan, D., Weinberg, R.A., The hallmarks of cancer, *Cell*, **2000**, 100 (1), 57-70
- <sup>10</sup> Hanahan, D., Weinberg, R.A., The hallmarks of cancer: the next generation, *Cell*, **2011**, 144 (5), 646-74
- <sup>11</sup> <http://med.stanford.edu/cutaneouslymphoma.html>
- <sup>12</sup> <https://www.cancer.gov/about-cancer/treatment/drugs/chlorambucil>
- <sup>13</sup> <https://journals.lww.com/oncology-times/blog/onlinefirst/pages/post.aspx?PostID=742>
- <sup>14</sup> Abeloff, M., *Clinical Oncology* (3<sup>rd</sup> edition), Churchill Livingstone, **2008**, 408-413
- <sup>15</sup> <https://www.cancernetwork.com/articles/principles-oncologic-pharmacotherapy>
- <sup>16</sup> Burns, C.M., Wortmann, R.L., Where will new drugs come from?, *The Lancet*, **2011**, 377, 97
- <sup>17</sup> <http://www.chemicalprobes.org/>
- <sup>18</sup> Goodsell, David S., The Molecular Perspective: The *ras* Oncogene, *Stem Cells*, **1999**, 17, 235-236
- <sup>19</sup> Cox, Adrienne D., Der, Channing J., Ras history: The saga continues, *Small GTPases*, **2010**, 1:1, 2-27
- <sup>20</sup> Malumbres, M., Barbacid M., RAS oncogenes: the first 30 years, *Nat Rev Cancer*, **2003**, 3, 459-465
- <sup>21</sup> Harvey J.J., An Unidentified Virus Which Causes the Rapid Production of Tumours in Mice, *Nature*; **1964**, 204, 1104-1105
- <sup>22</sup> Kirsten W.H., Mayer L.A., Morphologic responses to murine erythroblastosis virus, *J Natl Cancer Inst*; **1967**, 39, 311-335
- <sup>23</sup> Scolnick E.M., Rands E., Williams, D., Parks, W.P., Studies on the nucleic acid sequences of Kirsten sarcoma virus: a model for formation of a mammalian RNA containing sarcoma virus, *Journal of Virology*, **1973**, 12, 458-463
- <sup>24</sup> Scolnick E.M., Parks, W.P., Harvey sarcoma virus: a second murine type C sarcoma virus with rat genetic information, *Journal of Virology*, **1974**, 13, 1211-1219
- <sup>25</sup> Scolnick E.M., Papageorge, A.G., Shih, T.Y., Guanine nucleotide-binding activity as an assay for src protein of rat-derived murine sarcoma viruses, *Proc Natl Acad Sci USA*, **1979**, 76, 5355-5359



- <sup>26</sup> Shimizu, K., Goldfarb, M., Suard, Y., Perucho, M., Li, Y., Kamata, T. et al., Three human transforming genes are related to the viral ras oncogenes, *Proc Natl Acad Sci USA*, **1983**, 80, 2112-2116
- <sup>27</sup> Hall, A., Marshall, C.J., Spurr, N.K., Weiss, R.A., Identification of transforming gene in two human sarcoma cell lines as a new member of the ras gene family located on chromosome 1, *Nature*, **1983**, 303, 396-400
- <sup>28</sup> Gibbs, J.B., Sigal, I.S., Poe, M., Scolnick, E.M., Intrinsic GTPase activity distinguishes normal and oncogenic ras p21 molecules, *Proc Natl Acad Sci USA*, **1984**, 81, 5704-5708
- <sup>29</sup> Karnoub, A.E., Weinberg, R.A., Ras oncogenes: split personalities, *Nat Rev Mol Cell Biol*, **2008**, 9, 517-531
- <sup>30</sup> Tidyman, W.E., Rauen, K.A., The RASopathies: developmental syndromes of Ras/MAPK pathway dysregulation, *Curr Opin Genet Dev*, **2009**, 19, 230-236
- <sup>31</sup> Wennerberg, K., Rossman, K.L., Der, C.J., The Ras superfamily at a glance, *J Cell Sci*, **2005**, 118, 843-846
- <sup>32</sup> Colicelli, J., Human RAS superfamily proteins and related, *GTPases Sci STKE*, **2004**, RE13
- <sup>33</sup> Rossman, K.L., Der, C.J., Sondek, J., GEF means go: turning on RHO GTPases with guanine nucleotide exchange factors, *Nat Rev Mol Cell Biol*, **2005**, 6, 167-180
- <sup>34</sup> Aspuria, P.J., Tamanoi, F., The Rheb family of GTPbinding proteins, *Cell Signal*, **2004**, 16, 1105-1112
- <sup>35</sup> Bos, J.L., Fearon, E.R., Hamilton, S.R., Verlaan-de Vries, M., Van Boom, J.H., Van der Eb, A.J., Prevalence of ras gene mutations in human colorectal cancers, *Nature*, **1987**, 327, 293-297
- <sup>36</sup> Forrester, K., Almoguera, C., Han, K., Grizzle, W.E., Perucho, M., Detection of high incidence of K-ras Oncogenes during human colon tumorigenesis, *Nature*, **1987**, 327, 298-303
- <sup>37</sup> Rodenhuis, S., Van de Wetering, M.L., Mooi, W.J., Evers, S.G., Van Zandwijk, N., Bos, J.L., Mutational activation of the K-ras oncogene. A possible pathogenetic factor in adenocarcinoma of the lung, *N Engl. J. Med.*, **1987**, 317, 929-935
- <sup>38</sup> Almoguera, C., Shibata, D., Forrester, K., Martin, J., Arnheim, N., Perucho, M., Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes, *Cell*, **1988**, 53, 549-554
- <sup>39</sup> Sinn, E., Muller, W., Pattengale, P., Tepler, I., Wallace, R., Leder, P., Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: synergistic action of oncogenes in vivo, *Cell*, **1987**, 49, 465-475
- <sup>40</sup> Wood, L.D., Parsons, D.W., Jones, S., Lin, J., Sjoblom, T., Leary, R.J., The genomic landscapes of human breast and colorectal cancers, *Science*, **2007**, 318, 1108-1113
- <sup>41</sup> Sjoblom, T., Jones, S., Wood, L.D., Parsons, D.W., Lin, J., Barber, T.D., The consensus coding sequences of human breast and colorectal cancers, *Science*, **2006**, 314, 268-274
- <sup>42</sup> Jones, S., Zhang, X., Parsons, D.W., Lin, J.C., Leary, R.J., Angenendt, P., Core signalling pathways in human pancreatic cancers revealed by global genomic analyses, *Science*, **2008**, 321, 1801-1806
- <sup>43</sup> Parsons, D.W., Jones, S., Zhang, X., Lin, J.C., Leary, R.J., Angenendt, P., An integrated genomic analysis of human glioblastoma multiforme, *Science*, **2008**, 321, 1807-1812
- <sup>44</sup> Yan, H., Parsons, D.W., Jin, G., McLendon, R., Rasheed, B.A., Yuan, W., IDH1 and IDH2 mutations in gliomas, *N Engl J Med*, **2009**, 360, 765-773
- <sup>45</sup> Fearon, E.R., Vogelstein, B., A genetic model for colorectal tumorigenesis, *Cell*, **1990**, 61, 759-767
- <sup>46</sup> Shih, T.Y., Papageorge, A.G., Stokes, P.E., Weeks, M.O., Scolnick, E.M., Guanine nucleotide-binding and autophosphorylating activities associated with the p21src protein of Harvey murine sarcoma virus, *Nature*, **1980**, 287, 686-691

- 
- <sup>47</sup> Hurley, J.B., Simon, M.I., Teplow, D.B., Robishaw, J.D., Gilman, A.G., Homologies between signal transducing G proteins and *ras* gene products., *Science*, **1984**, 226, 860-862
- <sup>48</sup> McGrath, J.P., Capon, D.J., Goeddel, D.V., Levinson, A.D., Comparative biochemical properties of normal and activated human *ras* p21 protein, *Nature*, **1984**, 310, 644-649
- <sup>49</sup> Sweet, R.W., Yokoyama, S., Kamata, T., Feramisco, J.R., Rosenberg, M., Gross, M., The product of *ras* is a GTPase and the T24 oncogenic mutant is deficient in this activity., *Nature*, **1984**, 311, 273-275
- <sup>50</sup> Manne, V., Bekesi, E., Kung, H.F., H-ras proteins exhibit GTPase activity: point mutations that activate H-ras gene products result in decreased GTPase activity, *Proc Natl Acad Sci USA*, **1985**, 82, 376-380
- <sup>51</sup> Trahey, M., McCormick, F., A cytoplasmic protein stimulates normal N-ras p21 GTPase, but does not affect oncogenic mutants., *Science*, **1987**, 238, 542-545
- <sup>52</sup> Ballester, R., Marchuk, D., Boguski, M., Saulino, A., Letcher, R., Wigler, M., The NF1 locus encodes a protein functionally related to mammalian GAP and yeast IRA proteins, *Cell*, **1990**, 63, 851-859
- <sup>53</sup> Xu, G.F., Lin, B., Tanaka, K., Dunn, D., Wood, D., Gesteland, R., The catalytic domain of the neurofibromatosis type 1 gene product stimulates *ras* GTPase and complements *ira* mutants of *S. cerevisiae*., *Cell*, **1990**, 63, 835-841
- <sup>54</sup> Martin, G.A., Viskochil, D., Bollag, G., McCabe, P.C., Crosier, W.J., Haubruck, H., The GAP-related domain of the neurofibromatosis type 1 gene product interacts with *ras* p21., *Cell*, **1990**, 63, 843-849
- <sup>55</sup> Mitin, N., Rossman, K.L., Der, C.J., Signalling interplay in Ras superfamily function, *Curr Biol*, **2005**, 15, 563-574
- <sup>56</sup> Scheffzek, K., Lautwein, A., Kabsch, W., Ahmadian, M.R., Wittinghofer, A., Crystal structure of the GTPaseactivating domain of human p120GAP and implications for the interaction with Ras, *Nature*, **1996**, 384, 591-596
- <sup>57</sup> Kamata, T., Feramisco, J.R., Epidermal growth factor stimulates guanine nucleotide binding activity and phosphorylation of *ras* oncogene proteins, *Nature*, **1984**, 310, 147-150
- <sup>58</sup> Repasky, G.A., Chenette, E.J., Der, C.J., Renewing the conspiracy theory debate: does Raf function alone to mediate Ras oncogenesis?, *Trends Cell Biol*, **2004**, 14:639-47
- <sup>59</sup> Olivier, J.P., Raabe, T., Henkemeyer, M., Dickson, B., Mbamalu, G., Margolis, B., A Drosophila SH2-SH3 adaptor protein implicated in coupling the sevenless tyrosine kinase to an activator of Ras guanine nucleotide exchange, Sos, *Cell*, **1993**, 73, 179-191
- <sup>60</sup> Chardin, P., Camonis, J.H., Gale, N.W., Van Aelst, L., Schlessinger, J., Wigler, M.H., Human Sos1: a guanine nucleotide exchange factor for Ras that binds to GRB2, *Science*, **1993**, 260, 1338-1343
- <sup>61</sup> Quilliam, L.A., Huff, S.Y., Rabun, K.M., Wei, W., Park, W., Broek, D., Membrane-targeting potentiates guanine nucleotide exchange factor CDC25 and SOS1 activation of Ras transforming activity, *Proc Natl Acad Sci USA*, **1994**, 91, 8512-8516
- <sup>62</sup> Dickson, B., Sprenger, F., Morrison, D., Hafen, E., Raf functions downstream of Ras1 in the Sevenless signal transduction pathway, *Nature*, **1992**, 360, 600-603
- <sup>63</sup> Han, M., Golden, A., Han, Y., Sternberg, P.W., *C. elegans lin-45 raf* gene participates in let-60 ras-stimulated vulval differentiation, *Nature*, **1993**, 363, 133-140
- <sup>64</sup> Dent, P., Haser, W., Haystead, T.A., Vincent, L.A., Roberts, T.M., Sturgill, T.W., Activation of mitogen-activated protein kinase kinase by v-Raf in NIH 3T3 cells and in vitro, *Science*, **1992**, 257, 1404-1407

- 
- <sup>65</sup> Masuda, T., Kariya, K., Shinkai, M., Okada, T., Kataoka, T., Protein kinase Byr2 is a target of Ras1 in the fission yeast *Schizosaccharomyces pombe*, *J Biol Chem*, **1995**, 270, 1979-1982
- <sup>66</sup> Vojtek, A.B., Hollenberg, S.M., Cooper, J.A., Mammalian Ras interacts directly with the serine/threonine kinase Raf., *Cell*, **1993**, 74, 205-214
- <sup>67</sup> Rodriguez-Viciana, P., Warne, P.H., Dhand, R., Vanhaesebroeck, B., Gout, I., Fry, M.J., Phosphatidylinositol-3-OH kinase as a direct target of Ras, *Nature*, **1994**, 37, 527-532
- <sup>68</sup> Repasky, G.A., Chenette, E.J., Der, C.J., Renewing the conspiracy theory debate: does Raf function alone to mediate Ras oncogenesis?, *Trends Cell Biol*, **2004**, 14, 639-647
- <sup>69</sup> Spaargaren, M., Bischoff, J.R., Identification of the guanine nucleotide dissociation stimulator for Ral as a putative effector molecule of R-ras, H-ras, K-ras, and Rap, *Proc Natl Acad Sci USA*, **1994**, 91; 12609-12613
- <sup>70</sup> Hofer, F., Fields, S., Schneider, C., Martin, G.S., Activated Ras interacts with the Ral guanine nucleotide dissociation stimulator, *Proc Natl Acad Sci USA*, **1994**, 91, 11089-11093
- <sup>71</sup> Willumsen, B.M., Christensen, A., Hubbert, N.L., Papageorge, A.G., Lowy, D.R., The p21 ras C-terminus is required for transformation and membrane association., *Nature*, **1984**, 310, 583-586
- <sup>72</sup> Willingham, M.C., Pastan, I., Shih, T.Y., Scolnick, E.M., Localization of the src gene product of the Harvey strain of MSV to plasma membrane of transformed cells by electron microscopic immunocytochemistry, *Cell*, **1980**, 19, 1005-1014
- <sup>73</sup> Schafer, W.R., Kim, R., Sterne, R., Thorner, J., Kim, S.H., Rine, J., Genetic and pharmacological suppression of oncogenic mutations in *ras* genes of yeast and humans, *Science*, **1989**, 245, 379-385
- <sup>74</sup> Hancock, J.F., Magee, A.I., Childs, J.E., Marshall, C.J., All ras proteins are polyisoprenylated but only some are palmitoylated, *Cell*, **1989**, 57, 1167-1177
- <sup>75</sup> Choy, E., Chiu, V.K., Silletti, J., Feoktistov, M., Morimoto, T., Michaelson, D., Endomembrane trafficking of ras: the CAAX motif targets proteins to the ER and Golgi, *Cell*, **1999**, 98, 69-80
- <sup>76</sup> Vetter, I.R., Wittinghofer, A., The guanine nucleotide-binding switch in three dimensions. *Science*, **2001**, 294, 1299-1304
- <sup>77</sup> Cox, A. D., Der, C.J., Ras history: The saga continues, *Small GTPases*, **2010**, 1, 2-27
- <sup>78</sup> Saraste, M., Sibbald, P.R., Wittinghofer, A., The P-loop- a common motif in ATP-and GTP- binding proteins, *Trends Biochem. Sci.*, **1990**, 15, 430
- <sup>79</sup> Hillig, R.C., Structural and biochemical properties show ARL3-GDP as a distinct GTP binding protein, *Structure*, **2000**, 8, 1239
- <sup>80</sup> Roll-Mecak, A., Cao, C., Dever, T.E., Burley, S.K., X-Ray structures of the universal translational initiation factor IF2/eIF5B: conformational changes on GDP and GTP binding, *Cell*, **2000**, 103, 781
- <sup>81</sup> Prakash, B., Praefcke, G.J.K., Renault, L., Wittinghofer, A., Herrmann, C., Structure of human guanylate-binding protein 1 representing a unique class of GTP-binding proteins, *Nature*, **2000**, 403, 567
- <sup>82</sup> Lenzen, C., Cool, R.H., Prinz, H., Kuhlmann, J., Wittinghofer, A., Kinetic Analysis by Fluorescence of the Interaction between Ras and the Catalytic Domain of the Guanine Nucleotide Exchange Factor CDC25<sup>Mm</sup>, *Biochemistry*, **1998**, 19, 7420
- <sup>83</sup> Boriack-Sjodin, P.A., Mariana Margarit, S., Bar-Sagi, D., Kuriyan, J., The structural basis of the activation of Ras by Sos, *Nature*, **1998**, 394, 337-343

- 
- <sup>84</sup> Evelyn, C.R., Rational design of small molecule inhibitors targeting the Ras GEF, SOS1, *Chemistry and Biology*, **2014**, 21, 1618-1628
- <sup>85</sup> Sondermann, H., Structural Analysis of Autoinhibition in the Ras Activator Son of Sevenless, *Cell*, **2004**, 119, 393-405
- <sup>87</sup> Nassar, N., The 2.2 Å crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 complex with Rap1A and a GTP analogue, *Nature*, **1995**, 375-554
- <sup>88</sup> Rulack, T., Xia, F., Schlitter, J., Kotting, C., Gerwert, K., Ras and GTPase-activating protein (GAP) drive GTP into a precatalytic state as revealed by combining FTIR and biomolecular simulations, *PNAS*, **2012**, 109, 15295-15300
- <sup>89</sup> Rittinger, K., Walker, P.A., Eccleston, J.F., Smerdon, S. J., Gamblin, S.J., Structure at 1.65 Å of RhoA and its GTPase-activating protein in complex with a transition-state analogue., *Nature*, **1997**, 389, 758
- <sup>90</sup> Kraulis, P.J., *MOLSCRIPT*: a program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr.*, **1991**, 24, 946
- <sup>91</sup> Cox, A.D., Fesik, S.W., Kimmelman, A.C., Luo, J., Der, C.J., Drugging the undruggable RAS: Mission Possible?, *Nature*, **2014**, 13, 828-851
- <sup>92</sup> Grant, B. J., Novel allosteric sites on RAS for lead generation., *PLoS ONE*, **2011**, 6, e25711
- <sup>93</sup> Buhrman, G., Analysis of binding site hot spots on the surface of RAS GTPase., *J. Mol. Biol.*, **2011**, 413, 773–789
- <sup>94</sup> Taveras, A.G., RAS oncoprotein inhibitors: the discovery of potent, RAS nucleotide exchange inhibitors and the structural determination of a drug-protein complex, *Bioorg. Med. Chem.*, **1997**, 5, 125–133
- <sup>95</sup> Herrmann, C., Sulindac sulfide inhibits RAS signalling, *Oncogene*, **1998**, 17, 1769–1776
- <sup>96</sup> Gonzalez-Perez, V., Genetic and functional characterization of putative RAS/RAF interaction inhibitors in *C. elegans* and mammalian cells, *J. Mol. Signal.*, **2010**, 5, 2
- <sup>97</sup> Rosnizeck, I.C., Stabilizing a weak binding state for effectors in the human RAS protein by cyclen complexes, *Angew. Chem. Int. Ed Engl.*, **2010**, 49, 3830–3833
- <sup>98</sup> Margarit, S.M., Sonderman, H., Hall, B.E., Nagar, B., Hoelz, A., Pirruccello, M., Bar-Sagi, D., Kuriyan, J., Structural Evidence for Feedback Activation by Ras.GTP of the Ras-Specific Nucleotide Exchange Factor SOS, *Cell*, **2003**, 112, 685-695
- <sup>99</sup> Patgiri, A., Yadav, K.K., Arora, P.S., Bar-Sagi, D., An orthosteric inhibitor of the RAS–SOS interaction, *Nature Chem. Biol.*, **2011**, 7, 585–587
- <sup>100</sup> Maurer, T., Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity, *Proc. Natl. Acad. Sci. USA*, **2012**, 109, 5299–5304
- <sup>101</sup> Sun, Q., Discovery of small molecules that bind to KRAS and inhibit SOS-mediated activation, *Angew. Chem. Int. Ed. Engl.*, **2012**, 51, 6140–6143
- <sup>102</sup> Hilig, Roman C., Discovery of potent SOS1 inhibitors that blocks RAS activation via disruption of the RAS-SOS1 interaction, *PNAS*, **2019**, 116, 2551-2560
- <sup>103</sup> Shima, F., *In silico* discovery of small-molecule RAS inhibitors that display antitumor activity by blocking the RAS-effector interaction, *Proc. Natl Acad. Sci. USA*, **2013**, 110, 8182–8187
- <sup>104</sup> Ostrem, J. M., Peters, U., Sos, M.L., Wells, J.A., Shokat, K.M., KRAS (G12C) inhibitors allosterically control GTP affinity and effector interactions, *Nature*, **2013**, 503, 548–551

- 
- <sup>105</sup> Lim, S.M., Therapeutic targeting of oncogenic KRAS by a covalent catalytic site inhibitor, *Angew. Chem. Int. Ed Engl.*, **2014**, 53, 199–204
- <sup>106</sup> Papke, B., Der, C.J., Drugging RAS: Know the enemy, *Science*, **2017**, 355, 1158–1163
- <sup>107</sup> Lazo, J.S., Sharlow, E.R., Drugging undruggable molecular cancer targets, *Annu. Rev. Pharmacol. Toxicol.*, **2016**, 56, 23–40
- <sup>108</sup> Pylayeva-Gupta, Y., Grabocka, E., Bar-Sagi, D., RAS oncogenes: weaving a tumorigenic web, *Nat. Rev. Cancer*, **2011**, 11, 761–774
- <sup>109</sup> Bryant, K.L., Mancias, J.D., Kimmelman, A.C., Der, C.J. KRAS: feeding pancreatic cancer proliferation, *Trends Biochem. Sci.*, **2014**, 39, 91–100
- <sup>110</sup> Cox, A.D., Der, C.J., Philips, M.R., Targeting RAS membrane association: back to the future for anti-RAS drug discovery?, *Clin. Cancer Res.*, **2015**, 21, 1819–1827
- <sup>111</sup> Ahearn, I.M., Haigis, K., Bar-Sagi, D., Philips, M.R., Regulating the regulator: post-translational modification of RAS, *Nat. Rev. Mol. Cell Biol.*, **2011**, 13, 39–51
- <sup>112</sup> Schmick, M., K-Ras localizes to the plasma membrane by spatial cycles of solubilisation, trapping and vesicular transport, *Cell*, **2014**, 157, 459–471
- <sup>113</sup> Heidorn, S.J., Kinase-dead BRAF and oncogenic RAS cooperate to drive tumor progression through CRAF, *Cell*, **2010**, 140, 209–221
- <sup>114</sup> Peng, S.B., Inhibition of RAF isoforms and active dimers by LY3009120 leads to anti-tumor activities in RAS or BRAF mutant cancers, *Cancer Cell*, **2015**, 28, 384–398
- <sup>115</sup> Zhang, C., RAF inhibitors that evade paradoxical MAPK pathway activation., *Nature*, **2015**, 526, 583–586
- <sup>116</sup> Duncan, S., Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer, *Cell*, **2012**, 149, 307–321
- <sup>117</sup> Athuluri-Divakar, S.K., A small molecule RAS-mimetic disrupts RAS association with effector proteins to block signalling, *Cell*, **2016**, 165, 643–655
- <sup>118</sup> Ritt, D.A., Inhibition of Ras/Raf/MEK/ERK pathway signalling by a stress-induced phosphor-regulatory circuit, *Mol. Cell*, **2016**, 64, 875–887
- <sup>119</sup> Kaelin Jr. W.G., The concept of synthetic lethality in the context of anticancer therapy, *Nat. Rev. Cancer*, **2005**, 5, 689–698
- <sup>120</sup> Downward, J., RAS synthetic lethal screens revisited: still seeking the elusive prize, *Clin. Cancer Res.*, **2015**, 21, 1802–1809
- <sup>121</sup> Hanahan, D., Weinberg, R.A., Hallmarks of cancer: the next generation, *Cell*, **2011**, 144, 646–674
- <sup>122</sup> Maurer, T., Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity, *Proc. Natl. Acad. Sci. U.S.A.*, **2012**, 109, 5299–5304
- <sup>123</sup> Yang, S., Pancreatic cancers require autophagy for tumor growth, *Genes Dev.*, **2011**, 25, 717–729
- <sup>124</sup> Ying, H., Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism, *Cell*, **2012**, 149, 656–670
- <sup>125</sup> Son, J., Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway, *Nature*, **2013**, 496, 101–105

- <sup>126</sup> Winter, J.J., Anderson, M., Blades, K., Brassington, C., Breeze, A.L., Chresta, C., Embrey, K., Fairley, G., Faulder, P., Finlay, M.R., Kettle, J.G., Nowak, T., Overman, R., Patel, S.J., Perkins, P., Spadola, L., Tart, J., Tucker, J.A., Wrigley, G., Small Molecule Binding Sites on the Ras:SOS Complex Can Be Exploited for Inhibition of Ras Activation, *J. Med. Chem.*, **2015**, 58, 2265-2274
- <sup>127</sup> Ma, J., Cao, Q., McLeod, S.M., Ferguson, K., Gao, N., Breeze, A.L., Hu, J., Target-Based Whole-Cell Screening by (1) H NMR Spectroscopy, *Angew. Chem. Int. Ed. Engl.*, **2015**, 54, 4764-4767
- <sup>128</sup> Vo, U., Vajpai, N., Flavell, L., Bobby, R., Breeze, A.L., Embre, K., Golovanov, A.P., Monitoring Ras interactions with the nucleotide exchange factor Sos using site-specific NMR reporter signals and intrinsic fluorescence, *J. Biol. Chem.*, **2015**, jbc.M115.691238
- <sup>129</sup> Kyle, H.F., Wickson, K.F., Stott, J., Burslem, G.M., Breeze, A.L., Tomlinson, D.C., Warriner, S.L., Nelson, A., Wilson, A.J., Edwards, T.A., Exploration of the HIF-1 $\alpha$ /p300 binding interface using peptide and adhiron phage display technologies to locate binding hot-spots for inhibitor discovery, *Mol. BioSyst.*, **2015**, doi:10.1039/c5mb00284b
- <sup>130</sup> Fesik, S.W., Discovery of Aminopiperidine Indoles That Activate the Guanine Nucleotide Exchange Factor SOS1 and Modulate RAS Signaling, *J. Med. Chem.*, **2018**, 61, 6002-6017
- <sup>131</sup> Holdgate, G., Geschwindner, S., Breeze, A.L., Davies, G., Colclough, N., Temesi, D., Ward, L., Biophysical Methods in Drug Discovery from Small Molecule to Pharmaceutical Industry, *Protein-Ligand Interactions: Methods and Applications* (ed. Williams, M.A., Daviter, T.), *Methods Mol. Biol.*, **2013**, 1008, 327-355
- <sup>132</sup> Ishikawa, M., Hashimoto, Y., Improvement in Aqueous Solubility in Small Molecule Drug Discovery Programs by Disruption of Molecular Planarity and Symmetry, *J. Med. Chem.*, **2011**, 54, 1539-1554
- <sup>133</sup> Gough, D.J., Levy, D.E., Johnstone, R.W., Clarke, C.J., IFN $\gamma$  signalling-Does it mean JAK-STAT?, *Cytokine Growth Factor Rev.*, **2008**, 19, 383-394
- <sup>134</sup> Nguyen, H., Hiscott J., Pitha, P.M., The growing family of interferon regulatory factors, *Cytokine Growth Factor Rev.*, **1997**, 8, 293-312
- <sup>135</sup> Landis, S.H., Murray, T., Bolden, S., Wingo, P.A., Cancer statistics, *Cancer J. Clin.*, **1999**, 49, 8-31
- <sup>136</sup> Chung-Che, Chang, Jennifer, Lorek, Daniel, E. Sabath, Ying, Li, Christopher, R. Chitambar, Brent, Logan, Bal, Kampalath, Ronald, P. Cleveland, Expression of MUM1/IRF4 correlates with clinical outcome in patients with B-cell chronic lymphocytic leukaemia, *Blood Journal*, **2002**, 100, 4671-4675
- <sup>137</sup> Rasi, S., Gaidano, G., Shukla, V., Rinqink, L., IRF4 (interferon regulatory factor 4), *Atlas Genet. Cytogenet. Oncol. Haematol.*, **2014**
- <sup>138</sup> Carbone, A., Gloghini, A., Aldinucci, D., Expression pattern of MUM1/IRF4 in the spectrum of pathology of Hodgkin's disease, *British Journal of Haematology*, **2002**, 117, 366-372
- <sup>139</sup> Thuy, N. Do, Ucisik-Akkaya, E., Charronne, F.D., Morrison, B.A., Tevfik Dorak, M., An intronic polymorphism of IRF4 gene influences gene transcription *in vitro* and shows a risk association with childhood acute lymphoblastic leukemia in males, *Biochimica et Biophysica Acta*, **2010**, 1802, 292-300
- <sup>140</sup> Vipul, S., Runqing, L., IRF4 (interferon regulatory factor 4); *Atlas of Genetics and Cytogenetics in Oncology and Haematology*, **2014**, 18, 663-667
- <sup>141</sup> Yamagata, T., Nishida, J., Tanaka, S., Sakai, R., Mitani, K., A novel interferon regulatory factor family transcription factor, ICSAT/Pip/LSIRF, that negatively regulates the activity of interferon-regulated genes; *Mol Cell Biol.*, **1996**, 16, 1283-1294
- <sup>142</sup> Pham-Ledard, A., Prochazkova-Carlotti, M., Laharanne, E., Vergier, B., Jouary, T., IRF4 gene rearrangements define a subgroup of CD30-positive cutaneous T-cell lymphoma: a study of 54 cases, *J. Invest. Dermatol.*, **2009**, 130, 816-825

- 
- <sup>143</sup> Bisig, B., Gaulard, P., de Leval, L., New biomarkers in T-cell lymphomas, *Best Pract. Res. Clin. Haematol.*, **2012**, 25, 13–28
- <sup>144</sup> Aggarwal *et al.*, Crystal Structure of PU.1/IRF-4/DNA Ternary Complex, *Mol Cell*, **2002**, 10, 1097-1105
- <sup>145</sup> Koraki, R., Billeter, M., Wuthrich, K., MOLMOL: A program for display and analysis of macromolecular structures, *Journal of Molecular Graphics*, **1996**, 51-55
- <sup>146</sup> Nicholls, A., Sharp, K., Honig, B., Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins*, **1991**, 11, 281-296
- <sup>147</sup> Brivanlou, A.H., Darnell, J.E., Signal transduction and the control of gene expression, *Science*, **2002**, 94, 557-572
- <sup>148</sup> Abbott, A., On the offensive, *Nature*, **2002**, 416, 470-474
- <sup>149</sup> Heintel, D., Schreder, M., Strasser-Weippl, K., Vesely, M., Gisslinger, H., Drach, J., Gaiger, A., Jager, U., Ludwig, H., Expression of MUM1/IRF4 mRNA as a prognostic marker in patients with multiple myeloma, *Leukemia*, **2008**, 22, 441-445
- <sup>150</sup> Galustian, C. *et al.*, The anti-cancer agents lenalidomide and pomalidomide inhibit the proliferation and function of T regulatory cells, *Cancer Immunology Immunotherapy*, **2009**, 58, 1033-1045
- <sup>151</sup> Rajkumar, S.V. *et al.*, Lenalidomide plus high-dose dexamethasone versus lenalidomide plus low-dose dexamethasone as initial therapy for newly diagnosed multiple myeloma: an open-label randomised controlled trial, *The Lancet*, **2010**, 11, 29-37
- <sup>152</sup> Lopez-Girona, A. *et al.*, Lenalidomide downregulates the cell survival factor, interferon regulatory factor 4, providing a potential mechanistic link for predicting response, *British Journal of Haematology*, **2011**, 154, 325-336
- <sup>153</sup> Moros, A. *et al.*, Synergistic antitumour activity of lenalidomide with the BET bromodomain inhibitor CPI203 in bortezomib-resistant mantle cell lymphoma, *Leukemia*, **2014**, 28, 2049-2059
- <sup>154</sup> Remesh, S.G., Santosh, V., Escalante, C.R., Structural studies of IRF4 Reveal a Flexible Autoinhibitory Region and a Compact Linker Domain, *J. Biol. Chem.*, **2015**, 27779-27790
- <sup>155</sup> Mackerell, Alexander Jr., Zhang, H., Osterman, A., Kolhatkar, R., WO 2011006158, **2011**
- <sup>156</sup> Djaballah, H., Varmus, H. E., Shum, D., Somwar, R., Chucholowski, A., Thiruvazhi, M.S., WO 2008080056, **2008**
- <sup>157</sup> (a) Gabriel, S., Colman, J., Synthese des Pyridazins und seiner Derivate, *J. Chem. Ber.*, **1899**, 32, 398 (b) Maxwell, G.M., Ness, D., The effects of a new anti-hypertensive agent (endralazin) upon the general and coronary haemodynamics of the anaesthetized dog; *Eur. J. Pharmacol.*, **1981**, 8, 133-139 (c) Oates, H. F., Stoker, L. M., Studies in the rat on endralazine, a new antihypertensive drug structurally related to hydralazine, *Clin. Exp. Pharmacol. Physiol.*, **1981**, 8, 133–139 (d) Worms, P., Guedet, C., Biziere, K., Induction of turning by direct intrastriatal injection of dopaminomimetic drugs in mice: pharmacological analysis of a simple screening model, *Life Sci.*, **1996**, 39, 297-303
- <sup>158</sup> Nomoto, Y., Takai, H., Ohno, T., Nagashima, K., Ichimura, M., Mihara, A., Kase, H., Recent Developments in Pyridazine and Condensed Pyridazine Synthesis, *J. Med. Chem.*, **1996**, 39, 297-303
- <sup>159</sup> Xudong, C., Yin, C., Yifang, Z., Yu, L., Juencheng, Z., Xiangqing, X., Yinli, Q., Song, Z., Xin, L., Bi-Feng, L., Guisen, Z., *Journal of Medicinal Chemistry*, **2016**, 59, 2942–2961
- <sup>160</sup> Hegyi, G. *et al.*, Introduction to Practical Biochemistry, **2013**, Chapter 8, 8.7.1. Surface Plasmon Resonance
- <sup>161</sup> Swinney, D.C., Anthony, J., How were new medicines discovered?, *Nature*, **2011**, 10, 507-519

- <sup>162</sup> Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Delivery Rev.*, **1997**, 23, 3-25
- <sup>163</sup> Fesik, S.W., Discovering High-Affinity Ligands for Proteins: SAR by NMR, *Science*, **1996**, 274, 1531-1534
- <sup>164</sup> Congreve, M., Carr, R., Murray, C., Jhoti, H., A “rule of Three” for fragment-based lead discovery?, *Drug Discovery Today*, **2003**, 8, 876-877
- <sup>165</sup> Barker, J., Hestekamp, T., Whittaker, M., Integrating HTS and fragment-based drug discovery, *Drug Discov. World*, **2008**, 3, 69-75
- <sup>166</sup> Davies, T.G., Van Monfort, R.L.M., Williams, G., Jhoti, H., **2006**, Chapter 10: Pyramid: An integrated Platform for Fragment-based Drug Discovery, *Fragment Based Approaches in Drug Discovery* (193-214)
- <sup>167</sup> Hendrickson, W.A., Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of 3D structure, *EMBO J.*, **1990**, 9, 1665-1672
- <sup>168</sup> Chen, L. Q., Rose, J. P., Breslow, E., Yang, D., Chang, W. R., Furey, W. F., Sax, M., Wang, B.C., Crystal structure of a bovine neurophysin II dipeptide complex at 2.8 Å determined from the single-wavelength anomalous scattering signal of an incorporated iodine atom, *Proc. Natl Acad. Sci. USA*, **1991**, 88, 4240–4244
- <sup>169</sup> Dauter, Z., Novel approach to phasing proteins: derivatization by short cryosoaking with halides, *Acta Crystallographica*, **2000**, 56, 232-237
- <sup>170</sup> Cano, C. *et al.*, Recent advances in CDK inhibitors for cancer therapy, *Future Medicinal Chemistry*, **2018**, 10, 1-20
- <sup>171</sup> Schonbrunn, E. *et al.*, Discovery of a Potential Allosteric Ligand Binding Site in CDK2, *ACS chemical biology*, **2010**, 6, 492-501
- <sup>172</sup> Pearce, N. M., Krojer, T., Bradley, A. R., Collins, P., Nowak, R. P., Talon, R., Marsden, B. D., Kelm, S., Shi, J., Deane, C. M., Von Delft, F., A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density, *Nat. Commun*, **2017**, 8, 15123
- <sup>173</sup> Golding, B.T. *et al.*, Facilitation of displacements at the 6-position of purines by the use of 1,4-diazabicyclo[2.2.2]octane as leaving group, *J. Chem. Soc., Perkin Trans.*, 1, **1997**, 185
- <sup>174</sup> Lai, A.C.; Crews, C. M.; Induced protein degradation: an emerging drug discovery paradigm; *Nature Reviews Drug Discovery*, **2017**, 16, 101-114
- <sup>175</sup> Adjei, A.A., What is the right dose? The elusive optimal biologic dose in phase I clinical trials; *Journal of Clinical Oncology*, **2006**, 24, 4054-4055
- <sup>176</sup> Fedorov, Y. *et al.*, Off-Target effects by siRNA can induce toxic phenotype, *RNA*, **2006**, 12, 1188-1196
- <sup>177</sup> Jackson, A.L. *et al.*, Expression profiling reveals off-target gene regulation by RNAi, *Biotechnology*, **2003**, 21, 635-637
- <sup>178</sup> Qiu, S., Adema, C.M., Lane, T., A computational study of off-target effects of RNA interference, *Nucleic Acids Research*, **2005**, 33, 1834-1847
- <sup>179</sup> Doudna, J.A., Charpentier, E., The new frontier of genome engineering with CRISPR-Cas9, *Science*, **2014**, 346, 1258096
- <sup>180</sup> Hsu, P.D., Lander, E.S., Zhang, F., Development and applications of CRISPR-Cas9 for genome engineering, *Cell*, **2014**, 157, 1262-1278
- <sup>181</sup> Sander, J.D., Joung, J.K., CRISPR-Cas systems for editing, regulating and targeting genomes; *Nat. Biotechnology*, **2014**, 32, 347-355



- 
- <sup>182</sup> Crews, C.M., Targeting the undruggable proteome: the small molecules of my dreams, *Chem. Biol.*, **2010**, 17, 551-555
- <sup>183</sup> Overington, J.P., Al-Lazikani, B., Hopkins, A.L., How many drug targets are there?, *Nature*, **2006**, 5, 993-996
- <sup>184</sup> Campbell, J. *et al.*, Large-scale profiling of kinase dependences in cancer cell lines, *Cell. Rep.*, **2016**, 14, 2490-2501
- <sup>185</sup> Cowley, G.S. *et al.*, Parallel genome-scale loss of function screens in 216 cancer cell lines for the identification of context-specific genetic, *Sci. Data*, **2014**, 1, 140035
- <sup>186</sup> Wang, T. *et al.*, Identification and characterization of essential genes in the human genome, *Science*, **2015**, 350, 1096-1101
- <sup>187</sup> Hopkins, A. L., Groom, C. R., The druggable genome; *Nature*, **2002**, 1, 727-730
- <sup>188</sup> Duncan, J.S. *et al.*, Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer, *Nature*, **2012**, 149, 307-321
- <sup>189</sup> Hatzivassiliou, G. *et al.*, RAF inhibitors prime wild-type RAF to activate the MAPK pathway and enhance growth, *Nature*, **2010**, 464, 431-435
- <sup>190</sup> Heidorn, S.J. *et al.*, Kinase-dead BRAF and oncogenic RAS cooperate to drive tumor progression through CRAF, *Cell*, **2010**, 140, 209-221
- <sup>191</sup> Poulikakos, P.I., Zhang, C., Bollag, G., Shokat, K.M., Rosen, N., RAF inhibitors transactivate RAF dimers and ERK signalling in cells with wild-type BRAF, *Nature*, **2010**, 464, 427-430
- <sup>192</sup> Rauch, J., Volinsky, N., Romano, D., Kolch, W., The secret life of kinases: functions beyond catalysis, *Cell Commun. Signal.*, **2011**, 9, 23
- <sup>193</sup> Tan, X., Thapa, N., Sun, Y., Anderson, R.A., A kinase-independent role for EGF receptor in autophagy initiation, *Cell*, **2015**, 160, 145-160
- <sup>194</sup> Vivanco, I. *et al.*, A kinase-independent function of AKT promotes cancer cell survival, *eLife*, **3**, e03751
- <sup>195</sup> Weihua, Z. *et al.*, Survival of cancer cells is maintained by EGFR independent of its kinase activity, *Cancer Cell*, **2008**, 13, 385-393
- <sup>196</sup> Ciulli, A. *et al.*, Selective Small Molecule Induced Degradation of the BET Bromodomains, *ACS Chem. Biol.*, **2015**, 8, 1770-1777
- <sup>197</sup> Daniels, D.L. *et al.*, Quantitative Live-Cell Kinetic Degradation and Mechanistic Profiling of PROTAC Mode of Action., *ACS Chem. Biol.*, **2018**, 13, 2758-2770
- <sup>198</sup> Lecker, S.H., Goldberg, A.L., Mitch, W.E., Protein Degradation by the Ubiquitin-Proteasome Pathway in Normal and Disease States, *Journal of the American Society of Nephrology*, **2006**, 17, 1807-1819
- <sup>199</sup> Schulman, B.A., Harper, J.W., Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways, *Nature*, **2009**, 10, 319-331
- <sup>200</sup> Ye, Y., Rape, M., Building ubiquitin chains: E2 enzymes at work, *Nature*, **2009**, 10, 755-764
- <sup>201</sup> Deshaies, R.J., Joazeiro, C.A.P., RING domain E3 ubiquitin ligases, *Annu. Rev. Biochem.*, **2009**, 78, 399-434
- <sup>202</sup> Min, J.H. *et al.*, Structure of an HIF-1- $\alpha$ -pVHL complex: hydroxyproline recognition in signalling, *Science*, **2002**, 296, 1886-1889

- 
- <sup>203</sup> Winston, J.T. *et al.*, The SCF<sup>β-TRCP</sup>-ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in IκBα ubiquitination *in vitro*, **1999**, 13, 270-283
- <sup>204</sup> Petroski, M.D., Deshaies, R.J., Function and regulation of cullin-RING ubiquitin ligases, *Nature*, **2005**, 6, 9-20
- <sup>205</sup> Fang, S., Jensen, J.P., Ludwig, R.L., Voudsen, K.H., Weissman, A.M., Mdm2 is a RING finger-dependent ubiquitin protein ligase for itself and p53, *J. Biol. Chem.*, **2000**, 275, 8945-8951
- <sup>206</sup> Wade, M., Li, Y.C., Wahl, G.M., MDM2, MDMX and p53 in oncogenesis and cancer therapy, *Nature*, **2013**, 13, 83-96
- <sup>207</sup> Moulick, K. *et al.*, Affinity-based proteomics reveal cancer-specific networks coordinated by Hsp90, *Nature*, **2009**, 7, 818-826
- <sup>208</sup> Grenert, J.P. *et al.*, The amino-terminal domain of heat shock protein 90 (hsp90) that binds geldanamycin is an ATP/ADP switch domain that regulates hsp90 conformation, *J. Biol. Chem.*, **1997**, 272, 23843-23850
- <sup>209</sup> Stebbins, C.E. *et al.*, Crystal structure of an Hsp90-geldanamycin complex: targeting of a protein chaperone by an antitumor agent, *Cell*, **1997**, 90, 60-75
- <sup>210</sup> Prodromou, C. *et al.*, Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone, *Cell*, **1997**, 90, 60-75
- <sup>211</sup> Schneider, C. *et al.*, Pharmacological shifting of a balance between protein refolding and degradation mediated by Hsp90, *Proc. Natl. Acad. Sci.*, **1996**, 93, 14536-14541
- <sup>212</sup> Mimnaugh, E.G., Chavany, C., Neckers, L., Polyubiquitination and proteasomal degradation of the p185<sup>c-erbB-2</sup> receptor protein-tyrosine kinase induced by geldanamycin, *J. Biol. Chem.*, **1996**, 271, 22796-22801
- <sup>213</sup> Samuni, Y. *et al.*, Reactive oxygen species mediate hepatotoxicity induced by the Hsp90 inhibitor geldanamycin and its analogues, *Free Radic. Biol. Med.*, **2010**, 48, 1559-1563
- <sup>214</sup> Banerji, U. *et al.*, Phase I pharmacokinetic and pharmacodynamics study of 1,7-allylamino, 1,7-demethoxygeldanamycin in patients with malignancies, *J. Clin. Oncol.*, **2005**, 23, 4152-4161
- <sup>215</sup> Taldone, T., Gozman, A., Maharaj, R., Chiosis, G., Targeting Hsp90: small-molecule inhibitors and their clinical development, *Curr. Opin. Pharmacol.*, **2008**, 8, 370-374
- <sup>216</sup> Trepel, J., Mollapour, M., Giaccone, G., Neckers, L., Targeting the dynamic HSP90 complex in cancer.; *Nature*, **2010**, 10, 537-549
- <sup>217</sup> Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomised trials.; *Lancet*, **1998**, 351, 1451-1467
- <sup>218</sup> Jordan, V.C., Tamoxifen: A most unlikely pioneering medicine, *Nature*, **2003**, 2, 205-213
- <sup>219</sup> Wu, Y. *et al.*, Structural basis for an unexpected mode of SERM-mediated ER antagonism, *Mol. Cell*, **2005**, 18, 413-424
- <sup>220</sup> Vergote, I., Abraham, P., Fulvestrant, a new treatment option for advanced breast cancer: tolerability versus existing agents, *Ann. Oncol.*, **2006**, 17, 200-204
- <sup>221</sup> Weir, H.M. *et al.*, AZD9496: An oral estrogen receptor inhibitor that blocks the growth of ER-positive and *ESR1* mutant breast tumours in preclinical models, *Cancer Res.*, **2006**, <http://dx.doi.org/10.1158/0008-5472.can-15-2357>
- <sup>222</sup> Crawford, E.D. *et al.*, A controlled trial of leuprolide with and without flutamide in prostatic carcinoma, *N. Engl. J. Med.*, **1989**, 321, 419-424

- <sup>223</sup> Scher, H.I. *et al.*, Increased survival with enzalutamide in prostate cancer after chemotherapy, *N. Engl. J. Med.*, **2012**, 367, 1187-1197
- <sup>224</sup> Linja, M.J. *et al.*, Amplification and overexpression of androgen receptor gene in hormone-refractory prostate cancer, *Cancer Res.*, **2001**, 61, 3550-3555
- <sup>225</sup> <http://arvinas.com/category/press-releases/>
- <sup>226</sup> Neklesa, T.K., Crews, C.M., Chemical biology: greasy tags for protein removal, *Nature*, **2012**, 487, 308-309
- <sup>227</sup> Long, M.J., Gollapalli, D.R., Hedstrom, L., Inhibitor mediated protein degradation, *Chem. Biol.*, **2012**, 19, 629-637
- <sup>228</sup> Kubota, H., Quality control against misfolded proteins in the cytosol: a network for cell survival, *J. Biochem.*, **2016**, 291, 609-616
- <sup>229</sup> Neklesa, T. K. *et al.*, Small-molecule hydrophobic tagging-induced degradation of HaloTag fusion proteins., *Nat. Chem. Biol.*, **2011**, 7, 538-543
- <sup>230</sup> Teutsch, G. *et al.*, Non-steroidal antiandrogens: synthesis and biological profile of high-affinity ligands for the androgen receptor, *J. Steroid Biochem. Mol. Biol.*, **1994**, 48, 111-119
- <sup>231</sup> Gustafson, J. L. *et al.*, Small-molecule-mediated degradation of the androgen receptor through hydrophobic tagging, *Angew. Chem. Int. Ed.*, **2015**, 54, 9659-9662
- <sup>232</sup> Sakamoto, K. M. *et al.*, Protacs: chimeric molecules that target proteins to the Skp1-cullin-F-box complex for ubiquitination and degradation, *Proc. Natl Acad. Sci. USA*, **2001**, 98, 8554-8559
- <sup>233</sup> Griffith, E. C. *et al.*, Methionine aminopeptidase (type 2) is the common target for angiogenesis inhibitors AGM-1470 and ovalicin., *Chem. Biol.*, **1997**, 4, 461-471
- <sup>234</sup> Hon, W. C. *et al.*, Structural basis for the recognition of hydroxyproline in HIF-1 $\alpha$  by pVHL, *Nature*, **2002**, 417, 975-978
- <sup>235</sup> Min, J. H. *et al.*, Structure of a HIF-1 $\alpha$  -pVHL complex: hydroxyproline recognition in signalling, *Science* **2002**, 296, 1886-1889
- <sup>236</sup> Crews, C. *et al.*, Targeting the von Hippel-Lindau E3 Ubiquitin Ligase Using Small Molecules To Disrupt the VHL/HIF-1 $\alpha$  Interaction, *J. Am. Chem. Soc.*, **2012**, 134, 4465-4468
- <sup>237</sup> Schneekloth, J. S. Jr. *et al.*, Chemical genetic control of protein levels: selective *in vivo* targeted degradation, *J. Am. Chem. Soc.*, **2004**, 126, 3748-3754
- <sup>238</sup> Rodriguez-Gonzalez, A. *et al.*, Targeting steroid hormone receptors for ubiquitination and degradation in breast and prostate cancer, *Oncogene*, **2008**, 27, 7201-7211
- <sup>239</sup> Vassilev, L. T. *et al.*, *In vivo* activation of the p53 pathway by small-molecule antagonists of MDM2, *Science*, **2004**, 303, 844-848
- <sup>240</sup> Schneekloth, A. R.; Pucheault, M., Tae, H. S., Crews, C. M., Targeted intracellular protein degradation induced by a small molecule: en route to chemical proteomics, *Bioorg. Med. Chem. Lett.*, **2008**, 18, 5904-5908
- <sup>241</sup> Vu, B. *et al.*, Discovery of RG7112: a small-molecule MDM2 inhibitor in clinical development, *ACS Med. Chem. Lett.*, **2013**, 466-469
- <sup>242</sup> Itoh, Y., Ishikawa, M., Naito, M., Hashimoto, Y., Protein knockdown using methyl bestatin-ligand hybrid molecules: design and synthesis of inducers of ubiquitination-mediated degradation of cellular retinoic acid-binding proteins, *J. Am. Chem. Soc.*, **2010**, 132, 5820-5826

- <sup>243</sup> Buckley, D. L. *et al.*, HaloPROTACS: use of small molecule PROTACs to induce degradation of halotag fusion proteins, *ACS Chem. Biol.*, **2015**, 10, 1831–1837
- <sup>244</sup> Van Molle, I. *et al.*, Dissecting fragment-based lead discovery at the von Hippel–Lindau protein:hypoxia inducible factor 1 $\alpha$  protein–protein interface, *Chem. Biol.*, **2012**, 19, 1300–1312
- <sup>245</sup> Bondeson, D. P. *et al.*, Catalytic *in vivo* protein knockdown by small-molecule PROTACs, *Nat. Chem. Biol.*, **2015**, 11, 611–617
- <sup>246</sup> Ito, T. *et al.*, Identification of a primary target of thalidomide teratogenicity, *Science*, **2010**, 327, 1345–1350
- <sup>247</sup> Fischer, E. S. *et al.*, Structure of the DDB1-CRBN E3 ubiquitin ligase in complex with thalidomide, *Nature*, **2014**, 512, 49–53
- <sup>248</sup> Lu, J. *et al.*, Hijacking the E3 ubiquitin ligase cereblon to efficiently target BRD4, *Chem. Biol.*, **2015**, 22, 755–763
- <sup>249</sup> Crews, C. *et al.*, Cerehijacked with an “all small molecule” PROTAC; degrading the BET domain protein Brd4, *Chem Biol*, **2015**, 22, 755–763
- <sup>250</sup> Crews, C. *et al.*, First report of a small molecule VHL-based PROTAC; demonstrating the multi-turnover nature of PROTACs, *Nat. Chem. Biol.*, **2015**, 11, 611–617
- <sup>251</sup> Daniels, D.L. *et al.*, Quantitative Live-Cell Kinetic Degradation and Mechanistic Profiling of PROTAC Mode of Action, *ACS Chem. Bio.*, **2018**, 13, 2758–2770
- <sup>252</sup> Ciulli, A. *et al.*, Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4, *ACS Chem. Bio.*, **2015**, 10, 1770–1777
- <sup>253</sup> Ciulli, A. *et al.*, Structural basis of PROTAC cooperative recognition for selective protein degradation, *Nature Chem. Bio.*, **2017**, 13, 514–521
- <sup>254</sup> Kedra, C., Wehenkel, M., Choi, E., Han, H., Lee, H., Swanson, H., Kim, K., Impact of the linker length on the activity of PROTACs, *Mol. BioSyst.*, **2011**, 7, 359–364
- <sup>255</sup> Crews, C. *et al.*, Identification and Characterization of Von Hippel-Lindau-Recruiting Proteolysis Targeting Chimeras (PROTACs) of TANK-Binding Kinase 1, *J. Med. Chem.*, **2018**, 61, 583–598
- <sup>256</sup> Crews, C.M. *et al.*, Differential PROTAC substrate specificity dictated by orientation of recruited E3 ligase, *Nature Comm.*, **2019**, 10, 1–13
- <sup>257</sup> Lou, K.-J., PROTAC the protein, *SciBx*, **2010**, <http://dx.doi.org/10.1038/scibx.2012.514>
- <sup>258</sup> Bouchie, A., Allison, M., Webb, S., DeFrancesco, L., Nature Biotechnology’s academic spinouts of 2013, *Nat. Biotechnol.*, **2014**, 32, 229–238
- <sup>259</sup> <http://c4therapeutics.com/>
- <sup>260</sup> Chi, K. R., Drug developers delve into the cell’s trash-disposal machinery, *Nature*, **2016**, 15, 295–297
- <sup>261</sup> Wang, S. *et al.*, Discovery of MD-224 as a First-in-Class, Highly Potent, and Efficacious Proteolysis Targeting Chimera Murine Double Minute 2 Degradable Capable of Achieving Complete and Durable Tumor Regression, *J. Med. Chem.*, **2019**, 62, 448–466
- <sup>262</sup> Osborne, C. K., Wakeling, A., Nicholson, R. I., Fulvestrant: an oestrogen receptor antagonist with a novel mechanism of action, *Br. J. Cancer*, **2004**, 90, S2–S6
- <sup>263</sup> Yamada, A., Fujii, S., Mori, S., Kagechika, H., Design and Synthesis of 4-(4-Benzoylamino-phenoxy)phenol Derivatives As Androgen Receptor Antagonists, *ACS Med. Chem. Lett.*, **2013**, 4, 937–941

- <sup>264</sup> Marhefka, C.A., Wenqing, G., Chung, K., Kim, J., He, Y., Yin, D., Bohl, C., Dalton, J.T., Miller, D.D., Design, Synthesis, and Biological Characterization of Metabolically Stable Selective Androgen Receptor Modulators, *J. Med. Chem.*, **2004**, 47, 4, 993–998
- <sup>265</sup> Kussie, P.H., Svetlana, G., Marechal, V., Elenbaas, B. *et al.*, Structure of the MDM2 oncoprotein bound to the P53 tumor suppressor transactivation domain, *Science*, **1996**, 274, 948-953
- <sup>266</sup> Han, Kyou-Hoon *et al.*, Structural Details on mdm2-p53 interaction, *Journal of Biological Chemistry*, **2005**, 280, 38795–38802
- <sup>267</sup> Fuchs, S. Y., Adler, V., Buschmann, T., Wu, X.W., Ronai, Z., Mdm2 association with p53 targets its ubiquitination, *Oncogene*, **1998**, 17, 2543–2547
- <sup>268</sup> Wang, S. *et al.*, Temporal activation of p53 by a specific MDM2 inhibitor is selectively toxic to tumors and leads to complete tumor growth inhibition, *PNAS*, **2008**, 105, 3933-3938
- <sup>269</sup> Vassilev, L.T., Vu, B.T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C., Fotouhi, N., Liu, E.A., In vivo activation of the p53 pathway by small-molecule antagonists of MDM2, *Science*, **2004**, 303, 844-848
- <sup>270</sup> Buckley, C.C., Small-Molecule Control of Intracellular Protein Levels through Modulation of the Ubiquitin Proteasome System, *Angewandte Chemie International Edition*, **2014**, 53, 2312-2330
- <sup>271</sup> Burgess, A., Chia, K. M., Haupt, D., Thomas, D., Haupt, Y., Lim, E., Clinical overview of MDM2/X-targeted therapies, *Frontiers in Oncology*, **2016**, 6
- <sup>272</sup> Zhao, Y., Aguilar, A., Bernard, D., Wang, S., Small-Molecule Inhibitors of the MDM2–p53 Protein–Protein Interaction (MDM2 Inhibitors) in Clinical Trials for Cancer Treatment, *Journal of Medicinal Chemistry*, **2015**, 58, 1038-1052
- <sup>273</sup> Lv, P.-C., Sun, J., Zhu, H.-L., Recent advances of p53-MDM2 small molecule inhibitors (2011-present), *Current Medicinal Chemistry*, **2015**, 22, 618
- <sup>274</sup> Furet, P., Masuya, K., Kallen, J., Stachyra-Valat, T., Ruetz, S., Guagnano, V., Holzer, P., Mah, R., Stutz, S., Vaupel, A., Chene, P., Jeay, S., Schlapbach, A., Discovery of a novel class of highly potent inhibitors of the p53-MDM2 interaction by structure-based design starting from a conformational argument, *Bioorganic & Medicinal Chemistry Letters*, **2016**, 26, 4837-4841
- <sup>275</sup> Furet, P., Chène, P., De Pover, A., Stachyra-Valat, T., Hergovich Lisztwan, J., Kallen, J., Masuya, K., The central valine concept provides an entry in a new class of non peptide inhibitors of the p53-MDM2 interaction, *Bioorg. Med. Chem. Lett.*, **2012**, 22, 3498
- <sup>276</sup> Vaupel, A., Bold, G., De Pover, A., Stachyra-Valat, T., Hergovich Listzwan, J., Kallen, J., Masuya, K., Furet, P., Tetra-substituted imidazoles as a new class of inhibitors of the p53-MDM2 interaction, *Bioorg. Med. Chem. Lett.*, **2014**, 24, 2110
- <sup>277</sup> Gessier, F., Kallen, J., Jacoby, E., Chène, P., Stachyra-Valat, T., Ruetz, S., Jeay, S., Holzer, P., Masuya, K., Furet, P., Discovery of dihydroisoquinolinone derivatives as novel inhibitors of the p53-MDM2 interaction with a distinct binding mode, *Bioorg. Med. Chem. Lett.*, **2015**, 25, 3621
- <sup>278</sup> Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J., Pavletich, N. P., Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain, *Science*; **1996**, 274, 948
- <sup>279</sup> To avoid confusion the one letter code is used to name the amino acid residues of MDM2 while the three letter code is used for those of p53
- <sup>280</sup> Schneekloth, A. R., Pucheault, M., Tae, H. S., Crews, C. M., Targeted intracellular protein degradation induced by a small molecule: en route to chemical proteomics, *Bioorg. Med. Chem. Lett.*, **2008**, 18, 5904–5908

- <sup>281</sup> Antonarakis, E.S. *et al.*, Enzalutamide: the emperor of all anti-androgens, *Transl. Androl. Urol.*, **2013**, 2, 119-120
- <sup>282</sup> Magee, J.A. *et al.*, Direct, androgen receptor-mediated regulation of the FKBP5 gene *via* a distal enhancer element, *Endocrinology*, **2006**, 147, 590-598
- <sup>283</sup> Lunec, J. *et al.*, MYCN sensitizes neuroblastoma to the MDM2-p53 antagonists Nutlin-3 and MI-63, *Oncogene*, **2012**, 31, 752-763
- <sup>284</sup> Churcher, I., Protac-Induced Protein Degradation in Drug Discovery: Breaking the rules or Just Making New Ones?, *J. Med.Chem.*, **2018**, 61, 444-452
- <sup>285</sup> Nissan Chemical Ind., (2009). Pyridazinone derivatives as P2X7 receptor inhibitors and their preparation, pharmaceutical compositions and use in the treatment of rheumatoid arthritis. WO2009/057827 A1
- <sup>286</sup> Pu, Yu-Ming *et al.*, Efficient Copper-Catalyzed *N*-Arylation of Pyridazinones with Structurally Well-defined Copper Complex, *Tetrahedron Letters*, **2006**, 47, 149-153
- <sup>287</sup> Gillespie, Paul *et al.*, (2009). Preparation of benzylpiperazinepyrimidones as inhibitors of stearyl-CoA desaturase (SCD1). U.S. Pat. Appl. Publ., 20090149466
- <sup>288</sup> Gudmundsson, Kristjan S. *et al.* An improved large-scale synthesis of 2-amino-4-chloropyridine and its use for the convenient preparation of various polychlorinated 2-aminopyridines.; *Synthetic Communications*, **1997**, 27, 861-870
- <sup>289</sup> Leznoff, C. C., Svirskaya, P. I., Yedidia, V., Miller, J. M., *J. Heterocycl. Chem.*, **1985**, 22, 145
- <sup>290</sup> Arias, H. R. *et al.*, Novel Positive Allosteric Modulators of the Human  $\alpha 7$  Nicotinic Acetylcholine Receptor, *Biochemistry*, **2011**, 50, 5263-5278
- <sup>291</sup> Hamachi, I. *et al.*, LDAO-Based Chemical Labeling of Intact Membrane Proteins and Its Pulse-Chase Analysis under Live Cell Conditions, *Chemistry & Biology*., **2014**, 21, 1013-1022
- <sup>292</sup> Marzabadi, M.R. *et al.*, (2006). Preparation of aminothiazoles, aminotriazines, and aminobenzothiepinthiazoles as selective neuropeptide Y (NPY5) antagonist. U.S. Pat. Appl. Publ., 6989379
- <sup>293</sup> Meissner, A. *et al.*, Structure-activity relationships of 2-aminothiazoles effective against *Mycobacterium tuberculosis*, *Bioorganic & Medicinal Chemistry*, **2013**, 21, 6385-6397
- <sup>294</sup> Lawrence, Nicholas J., Pyridylthiazole-based ureas as inhibitors of Rho Associated protein kinases (ROCK1 and 2), *MedChemComm.*, **2012**, 3, 699
- <sup>295</sup> Kidwai, Mazaahir *et al.*, Eco-friendly synthesis of 2-aminothiazoles using Nafion-H as a recyclable catalyst in PEG-water solvent system, *Journal of Sulfur Chemistry*, **2011**, 32, 37-44
- <sup>296</sup> Ibrahim, S.A., Dyeing, Color Fastness and Antimicrobial Properties of Some Mono and Disazo Disperse Dyes Derived from Thiazole Moiety., *Fibres and Polymers*, **2013**, 14, 2061-2068
- <sup>297</sup> Stabile, R.G., Dicks, A.P., Microscale Synthesis and Spectroscopic Analysis of Flutamide, an Antiandrogen Prostate Cancer Drug., *J. Chem. Educ.*, **2003**, 80, 12, 1439
- <sup>298</sup> Gao, M *et al.*, Facile radiosynthesis of new carbon-11-labeled propanamide derivatives as selective androgen receptor modulator (SARM) radioligands for prostate cancer imaging, *Steroids*, **2011**, 76, 1505-1512
- <sup>299</sup> Ferla, Salvatore *et al.*, Design and synthesis of novel bicalutamide and enzalutamide derivatives as antiproliferative agents for the treatment of prostate cancer, *Journal of Med. Chem.*, **2016**, 118, 230-243

- 
- <sup>300</sup> Bertolo, Gabrielle *et al.*, A New Synthesis of 2,3-Dihydrobenzo[1,4]dioxin and 3,4-Dihydro-2H-benzo[1,4]oxazine Derivatives by Tandem Palladium-Catalyzed Oxidative Aminocarbonylation-Cyclization of 2-Prop-2-ynyloxyphenols and 2-Prop-2-ynyloxyanilines, *Journal of Organic Chem.*, **2006**, 71, 7895-7898
- <sup>301</sup> See reference 300
- <sup>302</sup> Loiseau, Francois A. *et al.*, Multigram Synthesis of Well-Defined Extended Bifunctional Polyethylene Glycol (PEG) Chains., *Journal of Med. Chem.*, **2004**, 69, 639-647
- <sup>303</sup> Amaral, S.P. *et al.*, Efficient Multigram Synthesis of the Repeating Unit of Gallic Acid-Triethylene Glycol Dendrimers., *Organic Letters.*, **2011**, 13, 4522-4525
- <sup>304</sup> Shan, Min *et al.*, Conformational Analysis of Bivalent Estrogen Receptor Ligands: From Intramolecular to Intermolecular Binding, *Chembiochem*, **2011**, 12, 2587-2598
- <sup>305</sup> Dorweiler, Jason D. *et al.*, Structural and NMR characterization of Sm (III), Eu (III), and Yb (III) Complexes of an amide based polydentate ligand exhibiting Paramagnetic Chemical Exchange Saturation Transfer abilities, *Inorganic Chem.*, **2009**, 48, 9365-9376
- <sup>306</sup> Wen, Wen-Hsien *et al.*, Synergistic Effect of Zanamivir-Porphyrin Conjugates on Inhibition of Neuraminidase and Inactivation of Influenza Virus., **2009**, 52, 4903-4910
- <sup>307</sup> Liu, Tu *et al.*, Coupling Across a DNA Helical Turn Yields a Hybrid DNA/Organic Catenane Doubly Tailed with Functional Termini, *J. Am. Chem. Soc.*, **2008**, 130, 10882-10883
- <sup>308</sup> Thomas, J. *et al.*, Size-specific ligands for RNA hairpin loops, *J. Am. Chem. Soc.*, **2005**, 127, 12434-12435